LUND UNIVERSITY

SCHOOL OF ECONOMICS AND MANAGEMENT

Is Sustainability Profitable?

The exploration of shareholder returns by applying an ESG strategy on the European market

Nils Betsholtz
Anton Lindström
Edvard Wennerberg

06/2020
Bachelor Thesis Accounting and Finance

Supervisors
Anna Glenngård
Gert Paulsson

Abstract

This paper examines the relationship between the ESG-score, including its pillars Environment, Social and Governance and market return from July 2002 through June 2018 by using the Stoxx Europe 600 index. The comparison is done by applying a portfolio approach and panel data fixed effect approach. By comparing some of the most used multifactor models, the CAPM and Fama-French three-factor model were found more suitable to price the sustainability sorted portfolios than the Carhart's four-factor model and Fama-French five-factor model. No significant difference in return between leading- and lagging sustainability corporations were found in the portfolio approach. However, some portfolios generated a positive abnormal return which may indicate a violation of the (semi-) efficient market hypothesis. The market is more (semi-) efficient after the financial crisis than before the financial crisis. In the panel data fixed effect approach, there was a negative relationship between ESG-score and market return. The negative relationship was higher after the financial crisis than before. Overall, the results were robust across the sustainability pillars in both the portfolio approach and panel data fixed effect approach. The findings suggest that corporations would not necessarily generate higher shareholder returns by undertaking ESG initiatives which may be suggested by legitimacy theory, although one can't exclude other non-financial gains. A negative relationship between shareholder returns and ESG would imply that the management has to balance the interests of the different stakeholders in their decision making.

Keywords: ESG, STOXX 600, Fama-French multifactor factor model, CAPM, Carhart four-factor model, Panel data fixed effect

Acknowledgement

We would like to thank Anna Glenngård and Gert Paulsson for their support and valuable advice which have increased the quality of this thesis.

Abbreviations

AIC	Akaike information criterion
BIC	Bayesian information criterion
BA	Big aggressive
BC	Big conservative
BH	Big high
BL	Big low
BLA	Big lagging
BLE	Big leading
BLUE	Best Linear Unbiased Estimator
BM	Big mid
BR	Big robust
BTM	Book to market
BV	Book value of equity
BW	Big weak
C4FM	Carhart four-factor model
CAPM	Capital asset pricing model
CLT	Central linear theorem
CMA	Conservative minus aggressive
CSP	Corporate social performance
CSR	Corporate social responsibility
DWS	Deutsche Gesellschaft für Wertpapiersparen (Fund manager)
EBT	Earnings before tax
EMH	Efficient market hypothesis
ENV	Environment
EOP	End of period
EPS	Earnings per share
ESG	Environment, Social and Governance
FF3FM	Fama-French three-factor model
FF5FM	Fama-French five-factor model
FRED	Federal Reserve Economic Data
GOV	Governance
HML	High minus low
KLD	Kinder, Lydenberg och Domini (sustainability index)

LIBOR	London Interbank Offered Rate
LMC	Logarithm of market capitalisation
LTA	Logarithm of total assets
MV	Market value of equity
NGO	Non-governmental organisation
OLS	Ordinary least square
OP	Operating profitability
PM	Profit Margin
PRI	Proponent of responsible investment
RMRF	Return of market portfolio minus risk-free rate
RMW	Robust minus weak
ROA	Return on assets
ROE	Return on equity
RSS	Residual sum of squares
SA	Small aggressive
SC	Small conservative
SH	Small high
SL	Small low
SLA	Small lagging
SLE	Small leading
SM	Small Mid
SMB	Small minus big
SOC	Social
SR	Small robust
STOXX600	Stoxx Europe 600 Index
STOXX1800	Stoxx 1800 Global Index
SW	Small weak
UMD	Up minus down
VIF	Variance inflation factor
VWR	Value weighted return

Figures, Equations and Tables

Figures

Fig. 1	Overview of previous studies
Fig. 2	ESG-score and its pillars
Fig. 3	Construction of portfolios for independent factors
Fig. 4	Construction of portfolios for dependent factors
Fig. 5	Number of corporations in the dependent portfolios

Equations

Eq. 1	Capital Assets Pricing Model
Eq. 2	Fama-French Five Factor Model
Eq. 3	Jensen's Alpha
Eq. 4	Total Return (security)
Eq. 5	Book-to-Market
Eq. 6	Investment Rate
Eq 7.1-4	Formulas for construction of the SMB factor
Eq 8.1-3	Formulas for construction of HML, CMA and RMW factors
Eq. 9	Sharpe Ratio
Eq. 10.1-2	Information Criteria
Eq. 11	Panel data fixed effect regression

Tables

Tbl. 1	The average monthly compounded portfolio excess returns
Tbl. 2	Portfolio approach – descriptive statistics of explanatory variables
Tbl. 3	Portfolio regressions for full period
Tbl. 4	Factor regressions based on FF5FM, C4FM and FF3FM
Tbl. 5	Model evaluation metrics
Tbl. 6	Sharpe ratios
Tbl. 7	Panel data fixed effect regressions

Table of Contents

1. Introduction 1
   1.1 Background 1
   1.2 Problem Identification 4
   1.3 Purpose of the Study 4
2. Theory 5
   2.1 Legitimacy Theory and Agency Theory 5
   2.2 Stakeholder Theory 6
   2.3 Efficient Market Hypothesis 7
   2.4 Capital Asset Pricing Model 9
   2.5 Multifactor Models developed by Fama-French and Carhart 10
   2.6 Jensen's Alpha 11
3. Literature Review 12
4. Methodology 15
   4.1 Research Strategy 15
   4.2 Data 16
   4.3 Input Data to Variable Construction 17
   4.4 Construction of Dependent- and Explanatory Variables 21
      4.4.1 Explanatory Variables 21
      4.4.2 Dependent Variables 22
      4.4.3 Essential consideration regarding the sorting Procedure 23
   4.5 Portfolio Evaluation and Model Comparison Metrics 24
      4.5.1 Sharpe Ratio 24
      4.5.2 Information Criteria 24
   4.6 Panel Data Regressions 25
   4.7 Data Tests 26
   4.8 Validity and Reliability 27
      4.8.1 Validity 27
      4.8.2 Reliability 27
5. Result and Analysis 29
   5.1 Descriptive Statistics and Data Tests 29
   5.2 Portfolio Approach 32
   5.3 Panel Data Approach 39
   5.4 Discussion of Theories 42
   5.5 Robustness Tests of Portfolio- and Panel Data Approach 43
   5.6 Comparison to Previous Research 44
6. Conclusion 46
7. References 47
8. Appendix 51

1. Introduction

1.1 Background

The importance of corporate social responsibility (CSR) has grown in recent years and is increasingly becoming a core aspect of corporate strategy, but has existed for a long time. The first signs of CSR initiatives are found in the annual reports of Shell in 1897, US Steel in 1905 and General Motors in 1916. Although these are early adopters, the prevalence of CSR reporting did not become common until the 1990s (Deegan and Unerman, 2011). In 1994, John Elkington established the concept of 3P which stands for people, planet and profit. This later evolved into the triple bottom line reporting, emphasising that an analysis of a corporation's performance has to consider, in addition to financial perspective, environmental and social perspectives to provide the users with a comprehensive representation of the organisation's corporate activities (Elkington, 1997). This led to the notion that not only management decisions regarding financial aspects but also those which impact the environment and social aspects have to be included in the analysis by investors (Deegan et al, 2011). Another evidence for this was put forward by Edmans (2011) who pointed to the fact that corporations included in the yearly rating of *the 100 best companies to work for in America,* which tends to focus on social aspects more than the average American corporation, also had a rate of return about 3,5% higher than the market average (Edmans, 2011).

As it became more evident that CSR activities were important to look at from an investor's point of view, the problem with there not being any clear definition to what can be included into the term became obvious. Consequently, the demand to easily rate and compare the performance of corporations within this field grew. In recent years, several financial data providers have begun to produce ESG-scores consisting of the three pillars Environment, Social and Governance. In short, the ESG-score is assigned to the corporations based on how they perform in the three different fields. It also makes it possible to quantify the information and establish relationships between the different scores and market performance.

International policymakers have had a key role in the emergence, design and implementation of ESG standards in corporate activities and reporting, for instance in creating efficiency, coordination as well as aligning the interest of organisations and its stakeholders. UN's Agenda 2030 consists of 17 sustainable development goals with 169 associated targets, which all 193 member states have signed (United Nations, 2015). Proponent of responsible investment (PRI) provides value to its signatory base by incorporating ESG aspects in investment and ownership decision making. PRI advocates long term orientation by trying to align conventional risk-adjusted profit maximisation with CSR (PRI Homepage, 2020). The EU is trying to achieve accountability, consistency and comparability in the non-financial reporting by stating guidelines on what aspects corporations should take into account and report (Council directive 2014/95/EU, 2014a). The mutual support between policymakers is clear as the EU refers to the UN global compact, among others, as a potential reporting system (Council directive 2014/95/EU, 2014b).

The issue of whether corporations should engage in CSR activities is tightly connected to the purpose of the corporation. This, together with the question of which specific stakeholders should be served by the corporation, has been debated widely since the 1960s. Milton Friedman (1970) stated that the responsibility of corporations is solely to generate returns to its shareholders, as this is the only responsibility of the management team as an agent. Friedman also argued that investment in CSR initiatives does not generate any shareholder value and therefore should not be undertaken by corporations. (Friedman, 1970). An opposing view was presented by Freeman (1984) who argued that a corporation has to devote efforts to satisfy all stakeholders in order to survive in the long term, especially in a turbulent and changing world. According to Freeman, corporations that actively work with CSR will therefore achieve higher profitability (Freeman, 1984).

Views and norms on which entities of society that should be served by the corporation vary greatly globally, and are determined on a national or local level. If corporations should address all stakeholders and thus engage in CSR activities, for whom should these activities be performed and what are corporations responsible for? These questions do not necessarily have clear answers. The social responsibilities of corporations are constructions of society in which corporations are

participants, which provides these corporations with the opportunity to shape what constitutes corporate social responsibility. (Borglund, De Geer and Sweet. 2012)

As corporations are devoting more efforts to reporting environmental, social and governance activities, the investment community has shown significant interest in exploring the relationships between financial or stock performance and ESG initiatives. Stevenson (2020) stated that it is frequently discussed if fund managers can outperform a traditional benchmark portfolio and thereby receive risk-adjusted return, captured by alpha, by just investing in stocks with a high ESG-score. In addition, Robin Braun, at DWS [a fund management company], pointed out that the advantage of lower volatility even if the portfolio is less diversified (Stevenson, 2020). In the past, there has mainly been an ongoing discussion of whether the market has financial metrics, e.g. profitability, which can be exploited by analysing financial data (market anomalies). Fama and French (1992) suggested in their three-factor model that market anomalies could be exploited by adding size and book-to-market in addition to the systematic risk factor. Small corporations and corporations with a high book-to-market value will outperform their large and low book-to-market counterparts. Hence, by simply looking at these two factors investors could expect a higher risk-adjusted return. In addition to the three-factor model, Carhart (1997) found evidence that momentum, a positive historical trend in the market price, influenced the future market price positively. Cochran and Wood (1984) highlighted risk as an important aspect because corporations with high risk tend to outperform the market during stock market upturns, while they perform poorly in stock market declines.

Later on, Fama and French (2015) added two more factors to their original three-factor model namely profitability and investment rate. Corporations with a high investment rate perform worse than corporations with low investment rate, and corporations with high profitability outperform those with low profitability. Stevenson (2020) concluded that there is hardly any conformity in the previous research enabling one to assume that a positive ESG screening strategy would contribute significantly to a positive alpha, as opposed to the previous discussion some research indicates that sin-stocks will outperform the market. Robert et. al. (2014) got a positive alpha whereas Galema et. al. (2008), Darren et. al (2012) and Yea Eun Jang (2019) received no significance. Franzen (2019) received a negative alpha. With regard to the financial parameters Darren et. al (2012) only

received lowered beta, whereas Galema et. al. (2008) just got a decreased book-to-market parameter.

1.2 Problem Identification

To our knowledge, there is no paper which has examined the relationship between ESG, all its pillars, and the shareholder return as well as comparing Capital Asset Pricing Model (CAPM), Fama and French three-factor model (FF3FM), Carhart four-factor model (C4FM) and Fama and French five-factor model (FF5FM). The effect of an ESG screening strategy on shareholder returns will also be examined by using a panel data fixed-effect model. By focusing on the Stoxx Europe 600 index (STOXX600) from July 2002 through June 2018 this thesis will also contribute to the current research as most studies have been performed on the American market. There is no conformity in the results of the previous research and this thesis will hence contribute by bringing clarity both regarding the selection of the most efficient multifactor model and the performance of an ESG screening strategy. The study has a broad range of practical implications and interests groups among others board of directors, fund managers, investors and policymakers.

1.3 Purpose of the Study

The purpose of this study is to separately explain the difference in return between high- and low ESG portfolios and its three pillars consisting of Environment-, Social- and Governance, by applying the CAPM, FF3FM, C4FM and FF5FM as well as a panel data fixed-effect model. The thesis will thereby be able to compare the explanatory power of different multifactor models and the performance of an ESG screening strategy.

2. Theory

2.1 Legitimacy Theory and Agency Theory

The legitimacy theory is based on the fact that corporations need to appear legitimate in the communities where they interact. Societies have norms and values that they find important, in addition to laws and regulations, and corporations are expected to follow these. These norms are not static but are constantly changing and therefore corporations have to adjust to them. According to the legitimacy theory, there is a social contract between organisations and the community in which they operate. This contract is difficult to define as it differs between interest groups and across time periods. Historically, the role of corporations has been seen as creating jobs for workers and goods/services for consumers. But as CSR has emerged, new expectations materialise around how corporations should act responsibly and take into account workers safety, the environment, consumer protection among many things. This is in addition to the laws and regulations which already govern corporate activities. According to the legitimacy theory, it becomes important for corporations to be perceived as being aligned with these values. Corporations which fail to do so will be penalized by consumer boycotts, actions by trade unions or difficulties to raise capital. (Deegan et al, 2011)

It can be difficult for individual stakeholders to gain insights into the operation of a corporation because of the information asymmetry that exists, information that the corporation itself releases about its impact on, for example, the environment, both social and natural, plays a major role in how the corporation is perceived. Since the corporation has more information than outside stakeholders, including the owners, it can influence how it wants to be viewed. It is important to remember that the way corporations appear could be more important than what it actually does.

The problem with information asymmetry is part of the bigger problem when one actor, the agent, makes decisions on behalf of somebody else, the principal. This is known as the agency dilemma. It arises from situations where the agents might be more motivated by self-interest than by the principal's interest. In the context of corporations, the management’s or agents, task is to make decisions, which benefits the shareholders, principals, and maximize their return of investment.

However, CEOs and board members might have different goals, for example making short term profits to receive bonuses and become more attractive for other corporations with more lucrative job offers. These goals might interfere with the long term goals of the shareholders (Jensen and Mechling, 1976). Because of asymmetric information, for instance, the agent has better knowledge of local markets, the principal can never be sure if the agent puts the principal's interests before his own. This is especially true when there is a conflict of interests. The bigger the information asymmetry the harder it is for the principal to control the agent. One way to decrease the gap is for the principal to monitor the agent. However, it requires costs for the principal to be fully aware of everything the agent does. If these monitoring costs are too high it is counter-effective to the purpose of hiring the agent. To circumvent the dilemma, contracts are often drawn as to have an alignment of goals such as commission pay and bonuses or by having third-party verifications of the agent's actions. This helps to avoid the dilemma and might lower the agency costs for the principal. The ESG-score is an example of verifications which makes it easier for stakeholders to gain knowledge of the operations of a corporation.

2.2 Stakeholder Theory

Stakeholder theory has many commonalities with legitimacy theory. However, while legitimacy theory treats the society as a whole and points out the importance of one common social contract, stakeholder theory, on the other hand, looks upon each individual group differently and highlights separate social contracts. Stakeholder theory is divided into two subcategories called the ethical branch and the managerial branch. (Deegan et al, 2011).

Within the ethical branch, sometimes referred to as the normative or moral, it is argued that the corporation should devote efforts to address all stakeholders regardless of their power. How much the organisation affects the lives of the stakeholder group determines the responsibilities the organisation has towards the group. The overall idea is that organisations should be managed for the benefit of all the stakeholders even if it means its financial performance will be worse, thus affecting shareholders negatively.

Edward Freeman (1984) defined a stakeholder as:
"Any identifiable group or individual who can affect the achievement of an organisation's objectives, or is affected by the achievement of an organisation's objectives." (Freeman, 1984, p.46)

This definition of stakeholders includes many groups, some which only marginally have any form of interaction with the organisation. Nevertheless, all have some basic rights which have to be acknowledged by the organisation. This includes both the responsibility to take action, or refrain from it, and to provide an account for the organisations' actions.

The managerial branch, the second branch of the stakeholder theory, looks at the stakeholders from the organisation's point of view. This means that the organisation will not view all the stakeholders as equals and therefore will not interact with them equally. Stakeholders will be assessed based on their power over the organisation and the more powerful a stakeholder is deemed to be, the more the organisation will have to focus on them. Mitchell, Agle and Wood (1997) expanded this idea and concluded there are three factors which stakeholders can be evaluated from. The three factors are: power, legitimacy and urgency and they decide which action the organisation needs to take. The more a stakeholder has of any or all three of these the more the corporation needs to take its interest into account.

In short, the more important a stakeholder group is to the goals of an organisation the more likely it is to be heard and have their interests taken into account. Different stakeholders may hold contradicting interests and the job of the management is to address them in accordance with how important they are to the organisation's goals (Mitchell et al, 1997).

2.3 Efficient Market Hypothesis

According to the efficient market hypothesis, an investor cannot achieve an abnormal return by just analyzing financial ratios and corporate characteristics since all public information is reflected in the stock price and available to all investors at a given point in time. The efficient-market hypothesis is incessantly debated and the theory is divided into three different degrees of efficiency:

The first one is *weak efficiency* in the market, the second is a *semi-efficient* market, and the third is *strong efficiency* in the market (Fama, 1970). Under weak market efficiency, the future price of stocks follows a random walk and cannot be predicted by conducting technical analysis. Consequently, it is not possible to identify predictive patterns based on historical price development. Investors using technical analysis are for instance trying to identify support- and resistance levels in the stock price, which are to be attributed to psychological reasons. The resistance level is an upper price level in which the stock price has a problem to exceed, whereas the support level is the corresponding lower price floor (Bodie, Kane and Marcus, 2014). Carhart (1997) is trying to exploit momentum in the development of the historical stock price that is identifying historical stock price trends which are thought to persist. Under weak market efficiency, fundamental analysis of corporate characteristics, as for instance income statement and the balance sheet, could in theory also result in information indicating that the stock is underpriced and thus enable an abnormal return (Bodie et al, 2014). If the market should instead be semi-efficient, one cannot predict price changes as all the public information is already included in the stock price. However, it would still be possible to predict the stock price if one has access to so-called insider information. (Fama, 1970)

This takes us to the third hypothesis, strong market efficiency, where even insider information is reflected in the stock price and an investor who has access to it cannot beat the market. The third hypothesis is impossible to test since there is no access to all insider information and thus cannot use it in an analysis. The laws against insider trading are supposed to prevent arbitrage profits on a weak- or semi-efficient market. (Byström, 2010). It is important to keep in mind that there will always be investors who outperform the market and want to attribute their gains to their own analysis of the market, but it can never be ruled out that this is solely due to chance. With many investors who daily trade in stocks, someone will always pick the right investment options by chance. According to the efficient market hypothesis, an investor in an at least semi-efficient market ought to have a weighted market portfolio in accordance with CAPM to maximize the risk-adjusted return. Nowadays, research indicates that capital markets are at least semi-efficient and that all available public information is reflected in the stock price (Byström, 2010). This means that the sustainability report will have an immediate and unpredictable impact on stock prices and consequently not enable any arbitrage profits.

2.4 Capital Asset Pricing Model

Capital Asset Pricing Model (CAPM) is a formula used to estimate the cost of capital or the expected return of a security. The model describes the relationship between expected return and systematic risk. The risk of a security can be either idiosyncratic or systematic. The idiosyncratic risk is firm-specific, which can be reduced by diversification, while the systematic risk is related to all securities. The CAPM-relationship is displayed by the following formula. (Berk and DeMarzo, 2014)

$$R_i = R_f + \beta_i\left(R_m - R_f\right) \tag{1}$$

The expected return of a security (R_i) consists of systematic risk and idiosyncratic risk, but the idiosyncratic risk can be eliminated through diversification. β_i determines the relationship between the price changes in the security *i* and an efficient portfolio. The risk premium for systematic risk is defined as the expected return of the market (R_m), less the return of a risk-free asset (R_f). According to the CAPM, a passive investment strategy holding a portfolio that only has exposure to systematic risk is the most efficient portfolio. (Berk and DeMarzo, 2014)

What the efficient portfolio is and what it contains has been discussed widely in finance. The portfolio can not include any firm-specific risk, meaning that the portfolio consists of a large set of assets and diversification cannot reduce any risk without reducing the return. The construction of the efficient portfolio is theoretically cumbersome. In doing so, the investor needs to know all expected return and volatility of all assets in the market. The CAPM solves this problem by showing that the market portfolio is the efficient portfolio if three assumptions are fulfilled. (Berk and DeMarzo, 2014)

1. *"Investors can buy and sell securities at competitive market prices (without incurring taxes or transaction costs) and can borrow and lend at the risk-free rate".*
2. *"Investors hold only efficient portfolios of traded securities - portfolios that yield the maximum expected return for a given level of volatility"*
3. *"Investors have homogeneous expectations regarding the volatilities, correlations and expected returns of securities"* (Berk and DeMarzo, 2014, p 379-380)

If investors are assumed to have homogeneous expectations, they hold the same efficient portfolios and thus own all stocks in the market, the market portfolio has to represent the efficient portfolio (Berk and DeMarzo, 2014). The assumptions underlying CAPM are very strong and not always describe investor behaviour. Therefore, the investor could ask whether the market portfolio is efficient.

2.5 Multifactor Models developed by Fama-French and Carhart

Fama and French (1992) expanded on the CAPM by introducing FF3FM to better explain asset pricing. Besides the market factor in CAPM already discussed above, Fama and French (1992) also included two additional factors capturing size, *market capitalisation*, and value, *book-to-market*. The reason for adding the two additional factors of size and value is because small and low valued stocks tend to outperform larger and highly valued peers. (Fama and French, 1992). The size effect is explained by Fama and French as smaller corporations have more room to grow. Small corporations also often encounter funding problems which once solved can spur growth (Fama and French 2015). Another expansion to FF3FM was proposed by Michael Carhart who added a momentum factor (Carhart 1997), known as the Carhart four-factor model (C4FM).

Fama and French (2015) expanded again on their three-factor model and added two more factors which were based on operating profitability (OP) and investment rate (IR). This is known as the Fama-French five-factor model (FF5FM). Stocks associated with higher future earnings tend to have higher returns and Fama and French used current earnings as a proxy for future earnings. With regards to investments, corporations that invest intensely tend to have lower stock returns (Fama and French 2015). The FF5FM is presented below:

$$R_{pt} - R_{ft} = \alpha_{pt} + \beta_{pt}(R_{mt} - R_{ft}) + s_{pt}SMB + m_{pt}HML + r_{pt}RMW + c_{pt}CMA + e_{pt} \quad (2)$$

The dependent variable R_{pt} - R_{ft} is the portfolio excess return and $R_{mt} - R_{ft}$ is the market excess return. SMB stands for small minus big and it accounts for the size effect. The other three factors stand for high minus low (*HML*), robust minus weak (*RMW*) and conservative minus aggressive (*CMA*) and take into account the value-, profitability- and investments effect respectively. The

five-factor model is estimated to explain about 71 - 94 % of the returns on the stocks. (Fama and French, 2015)

2.6 Jensen's Alpha

Jensen´s alpha is a measurement to evaluate the risk-adjusted return of a portfolio, and it is defined as the actual return subtracted by the expected return of the benchmark portfolio. Under the semi-strong efficient market hypothesis, the benchmark portfolio is only based on the market risk factor of the CAPM. The difference between the actual- and expected return shows how much of the portfolio return that can be accredited to the investor's ability versus the overall market return, and other risk factors, such as HML, in multifactor models. Jensen's alpha of an investment is positive if one succeeds to deliver a positive risk-adjusted return and vice versa. (Jensen, 1967)

$$\alpha_{it} = R_{it} - \left(R_{ft} + \beta_{it}\left(R_{mt} - R_{ft}\right)\right) \tag{3}$$

The formula above describes Jensen's alpha of the portfolio (α_{pt}), the return on the applied portfolio (R_{pt}), the return of the benchmark portfolio (R_{mt}), R_{ft} is the risk-free return and β_{pt}, beta, is the mark risk exposure. Jensen's alpha is used as a performance evaluation measurement of fund managers, or other investors and measures the return in comparison to a given benchmark portfolio, originally the CAPM.

3. Literature Review

ESG-score and its impact on the corporation's financial performance has been researched extensively, but its relation to shareholder returns has not been examined to the same degree. The reader should note that previous research focuses on different aspects of the corporation. While some studies focus on the relationship between sustainability and accounting-based metrics such as return on assets (ROA) or profitability, others focus on the total shareholder return, change in stock price plus dividends/buybacks. One may on average expect a positive relationship between financial performance and market performance. Hence, financial metrics are frequently used in multifactor models to explain market performance.

The connection between ESG-score and the corporation's financial performance has been examined by Cochran and Wood, who found a weak positive correlation between CSR work and financial profitability (Cochran and Wood, 1984). Other more contemporary studies have also found weak positive links between CSR work and financial metrics such as Minutolo, Kristjanpoller and Stakeley from 2019. However, no consensus exists and other scholars have concluded that no correlation can be found (Landi and Sciarelli, 2019). Barnett and Salomon (2011) showed that the correlation between CSR and financial performance is U-shaped. Corporations that demonstrate mediocre CSR work perform worse, while those who refrain from doing CSR and those who excel can demonstrate higher profitability. This is likely due to the fact that corporations with mediocre CSR work have costs for their CSR work but it is not good enough to have an impact on stakeholders such as consumers. (Barnett and Solomon, 2011).

Studies with a focus on the relationship between total shareholder return and ESG include Robert et. al. (2014) who matched high- and low ESG American corporations based on total assets, return on assets (ROA), return on equity (ROE), leverage, turnover and market-to-book ratio in order to make sure that the corporate characteristics were not significantly different from each other. The market return of corporations with high- and low ESG-scores was studied on the USA market between 1993 and 2010 using the C4FM. Corporations with high ESG-scores were found to significantly outperform corporations with low ESG-scores. In addition, both high- and low ESG corporations significantly outperformed the market as Jensen's alpha, in all regressions, were

significantly positive different from zero, although the difference at least partly can be attributed to survivorship bias.

Darren et. al (2012) examined the relationship between corporate social performance, CSR, which was proxied by the ESG-score and market value performance of US corporations. The study examined the period between 1998 and 2007 by using the C4FM. The industries were divided into leading- and lagging CSP industries and subsequently the corporations were sorted within leading and lagging industries based on CSP. The study was not able to find any significant difference in performance with regard to size, book-to-market and momentum, while the high-ranked CSP corporations were exposed to lower systematic risk. Material CSP differences were considered incorporated in the market price, resulting in no significant difference in alpha and thereby supporting the efficient market hypothesis.

Galema et. al. (2008) found that the effect of social responsible investments (SRI), using the KLD score of US corporations, goes through a lowered book-to-market rather than through a higher alpha. It could be interpreted as a reflection of the demand difference between high- and low CSP corporations. According to the authors, the majority of the research has not been able to establish a link between the CSP and alpha. There could also potentially be a difference of the effect within the different dimensions of a sustainability score although no evidence was found for this hypothesis.

Franzen (2019) studied the impact of ESG-score on market performance in the American market, S&P 500, between 2002 and 2017 by using the FF5FM. The results showed that low ESG-score portfolios outperformed high ESG-score portfolios both with regard to Jensen's alpha and the risk-adjusted return measured by Sharpe ratio. The panel data fixed effect regression generated mainly insignificant results. Yea Eun Jang (2019) used the C4FM in the European market, Stoxx Europe 600 index, between 2002 and 2018 by applying a regression with the difference in return between portfolios with a high ESG-score and portfolios with a low ESG-score. The result gave no significance at the explanatory factors, whereas Jensen's alpha turned out to indicate a lower return using a positive ESG screening strategy. The significant negative returns were confirmed in the pooled regression approach.

Accounting based financial metrics and ESG						
Author	**Year**	**Time period**	**Geography**	**ESG data provider (ESG measure)**	**Financial metrics (FM)**	**Relation between CSR and FM**
Barnet et al.	2011	1998 - 2006	USA	KLD Research and Analytics (ESG-score)	ROA Net income	Positive / Negative
Cochran et al.	1984	1970 - 1979	USA	Reputation index/ content analysis	OE / Assets OE/Sales EV	Positive
Landi et al.	2019	2007 - 2015	Italy	Standard Ethics Agency (ESG-score)	Stock return	No significant
Minutolo et al.	2019	2009 - 2015	USA	Bloomberg (ESG-score)	Tobin's Q ROA	Positive
Shareholder return and ESG						
Author	**Year**	**Time period**	**Geography**	**ESG data provider (ESG measure)**	**Methodology**	**Relation between ESG and Shareholder Return**
Darren et al.	2012	1998 - 2007	USA	RobecoSAM (CSA Score)	C4FM	No significant
Franzen	2019	2002 - 2017	USA	Datastream/Eikon (Asset 4 ESG)	FF5FM/Panel data fixed effect	Negative/ No significance
Galema et al.	2008	1992 - 2006	USA	KLD Research and Analytics (ESG-score)	Book-to-market regression and Fama-Macbeth Regression	No significant
Robert et al.	2014	1993 - 2010	USA	RobecoSAM (CSA Score)	C4FM	Positive
Yea Eun Jang	2019	2002 - 2018	Europe	Datastream/Eikon (Asset 4 ESG)	C4FM/Pooled Regression/ Cross-sectional regression	Negative on all methodologies

Figure 1: Overview of previous studies.

4. Methodology

4.1 Research Strategy

This study has been based on a deductive process, where previously discussed theories have been applied to explain the return of an ESG screening strategy. A quantitative method has been used where second-hand data was collected from Bloomberg and Eikon. The data has been used to generate regression results which either confirm or reject the theories (Bryman and Bell, 2017). In recent decades, the growth of awareness regarding sustainability issues has changed the norms and values of societies. In accordance with legitimacy and stakeholder theory, it is hence likely that corporations have to adjust to the expectations of the public to attain a good market performance. Legitimacy and stakeholder theory have been selected because they try to explain why corporations invest in sustainability related issues as well as the effects from the sustainability investments. The efficient market hypothesis has been used because the theory shows how and why sustainability activities affect the price of stocks.

The ESG screening strategy has been evaluated by applying a portfolio approach and panel data fixed effect approach. The portfolio approach was a time series regression of each size-sustainability sorted portfolio, where the model specification was based on Fama and French (1993) and Fama and French (2015). In the panel data approach, fixed-effect has been used to control for time-invariant attributes in each individual and time-specific effects were captured by using dummy variables for each year. One year has been excluded and was thereby used as a benchmark to the other years in order to avoid the dummy variable trap. Under the dummy variable trap, the dummies will always add one to the intercept causing perfect multicollinearity between the intercept vector and the combined vector of all dummy variables (Brooks, 2014). Robustness tests have been conducted before and after the financial crisis. The financial crisis period is defined in line with Lins et.al (2017), namely from August 2008 through March 2009.

4.2 Data

Data has been retrieved from the Eikon, Bloomberg and Federal Reserve Economic Data (FRED) for the period from July 2002 through June 2018. The time period has been chosen as the ESG-score was first published by Eikon in 2002. STOXX600 has been used as a proxy for the European market in the formation of the portfolios. The index includes securities from 17 countries in Europe and around 95% of the float market capitalisation of the 17 European countries (STOXX Homepage, 2020). The composition of the STOXX600 has been updated on an annual basis based on the historical composition at the end of June each year between 2002 and 2018. This avoids the survivorship bias which means the corporations observed at a given point in time likely have performed better than delisted corporations and hence may share other corporate characteristics such as ESG-score. Monthly return data based on the constructed portfolios has been compiled from July 2002 through June 2018 which provides the study with 192 months of observations. Interest rates of the 1-month London Interbank Offered Rate (LIBOR) have been collected from FRED and transformed into an effective monthly interest rate. The panel data fixed effect regression has been based on the corporations of STOXX600, but the number of corporations varies due to listing and delisting from the index during the period from July 2002 through June 2018.

Some securities in the STOXX600 have been excluded due to either financial industry classification or lack of data. Financial corporations have been excluded as they tend to have a higher leverage level due to their business model rather than as an indication of financial distress, as would have been the case for non-financial corporations (Fama and French, 1993). If financial corporations would have been included, they would have had a significant impact on the sample and the result would hence be non-representative. Securities have also been excluded if data cannot be retrieved to enable the construction of risk mimicking portfolios, for instance, market capitalisation was required to construct SMB, see section 4.3. Exclusions have also been made if the return data was missing for at least 3 out of 12 months in a given year. The number of excluded corporations due to missing data has significantly decreased during the studied period.

4.3 Input Data to Variable Construction

Risk-free Rate

To proxy the risk-free interest rate, the 1-month London Interbank Offered Rate (LIBOR) has been used. Fama and French (1993) used the one-month US treasury bill rate in their study on the American market.

Total Return - Security Level

The calculation of monthly return of every security accounts for movements in the stock price as well as for shareholder distribution initiatives such as dividends or stock buybacks (represented by Div_t).

$$Total\ Return = \frac{Share\ price_t - Share\ price_{t-1}}{Share\ price_{t-1}} + \frac{Div_t}{Share\ price_{t-1}} \quad (4)$$

Market Index

As a proxy for the market mimicking portfolio, Fama and French (1993) used the value-weighted excess return of the six portfolios constructed based on size and book-to-market, which means that they integrated all the listed corporations on NYSE-, Amex- and NASDAQ indexes. As discussed by Berk and DeMarzo (2014), it is practically impossible to construct a value-weighted market portfolio including all securities across assets, therefore one has to use a proxy for the market index portfolio. In this thesis, the excess return of the Stoxx 1800 Global Index (STOXX1800) has been used as it covers a broad range of corporations from the major markets in the world.

Size

In accordance with Fama and French (1993), size has been proxied by market capitalisation of the corporation on an annual basis at the end of June. Securities that had a larger market capitalisation than the median corporation in the STOXX600 were defined as *Big*, and securities with a smaller market capitalisation than the median corporation were defined as *Small*. The data has been collected at the end of June annually.

Book-to-Market

Book-to-market (BTM) value has been calculated at the end of June year t and is defined as the book value of common equity (BV) at the end of the fiscal year t-1 divided by the market value of equity (MV) at end of the fiscal year t-1. Book value of common equity was in turn calculated as the total assets at the end of fiscal year t-1 subtracted by the total liabilities at end of the fiscal year t-1. The market value of equity is defined as the total number of shares outstanding at the end of fiscal year t-1 multiplied by the market price per share at the same date (Fama and French, 1993). The top 30% of companies with the highest BTM were classified as *High* and the lowest 30% were classified as *Low*.

$$BTM_t = \frac{BV_{t-1}}{MV_{t-1}} \tag{5}$$

Momentum

Based on the Carhart (1997), the momentum factor was constructed based on the total shareholders return of the previous twelve months. The portfolios have been rebalanced on an annual basis at the end of June year t based on the total return of the previous twelve months. The *Up minus Down* (UMD) portfolio consists of the 30% highest-performing stocks minus the 30% lowest-performing stocks.

Profitability

In Fama and French (2015), the operating profitability (OP) was calculated by taking the annual revenues subtracted by the cost of goods sold, interest expenses and administrative-, general-, selling- expenses divided by the book value of equity. All the data has been collected at the end of the fiscal year t-1. In this thesis, earnings before tax have been used, meaning that interest income was included in the profitability metric as opposed to Fama and French (2015). It was not possible to use the income statement post which corresponded to Fama and French (2015), as Eikon did not provide data for a sufficient number of corporations. Note that interest income will in most cases have an insignificant influence on the result, as non-financial corporations do not tend to accumulate cash or financial instruments as this is not part of the core business and could represent a non-efficient asset allocation practice.

Securities have been divided into three categories with regards to their profitability, which were used to construct the factor mimicking portfolios. Corporations with profitability over the 70th

percentile were classified as *Robust* and corporations below the 30th percentile were classified as *Weak.* All other stocks in between the 30th and 70th percentile were classified as *Middle* (Fama and French, 2015).

Investment Rate

The investment rate was used as a proxy to capture the tendency of corporations to invest in their business. The rate is defined as:

$$Investment\ Rate = \frac{Total\ Assets_{t-1} - Total\ Assets_{t-2}}{Total\ Assets_{t-2}} \quad (6)$$

Securities have been divided into three categories with regards to their profitability for the creation of risk mimicking portfolios. Corporations with investment rates over the 70th percentile were classified as *Aggressive* and below the 30th percentile as *Conservative*. All other stocks were classified as *Middle*. (Fama and French, 2015)

ESG-score

Data on the ESG-score and its pillars Environment, Social and Governance are provided by Thomson Reuters (Thomson Reuters, 2018). The scores are provided both as a percentile (0-100) and as a grade with twelve categories ranging from D- to A+. This study has used the scores because the data has a higher classification (ratio scale). The ESG-score is calculated by weighting the subcategories seen in the figure below.

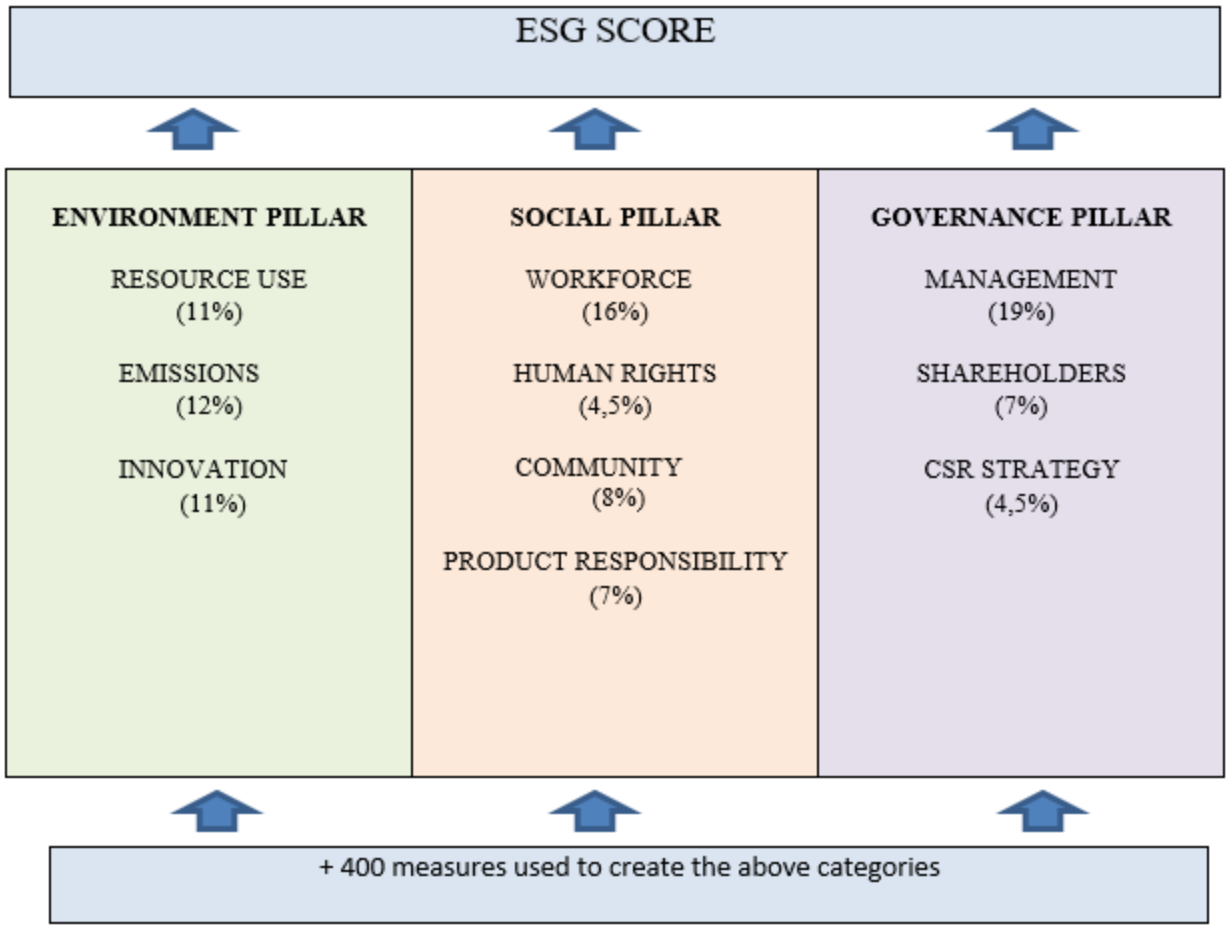


Figure 2 - The figure describes the breakdown of the ESG-score and the weight of the various components.

As indicated in the table above, the three ESG pillars environment, social and governance are in turn based on more than 400 indicators. The data is originally derived from financial- and sustainability reports as well as from media coverage and NGOs. In addition to the original ESG-score, there is an ESG *combined* score which is based on the original ESG-score but adds a dimension called the *controversies category*. The ESG *combined* score consists of an additional 23 measurements which include lawsuits or controversies reported in the news (Thomson Reuters, 2018). The traditional ESG-score has been used in this study as the ESG *combined* score is a relatively new rating, which thereby has limited data.

4.4 Construction of Dependent- and Explanatory Variables

4.4.1 Explanatory Variables

The excess return on the market risk factor has been calculated by taking the value-weighted return (VWR) of a market mimicking portfolio and subtracting the risk-free rate. In the construction of the risk mimicking portfolios used to calculate the risk factors, the corporations, in STOXX600, have first been split at the median with regard to size. Afterwards, book-to-market, profitability and investment rate have been independently and separately split into three groups based on the 30th and 70th percentile. The intersection between size and each of book-to-market, profitability and investment rate has been used to construct the factor mimicking portfolios (Fama and French, 1993). Six 2*3 sorted portfolios were constructed for each combination of size and book-to-market, investment and profitability respectively, giving 3 sets of portfolios with 6 portfolios in each, thereby 18 portfolios in total (Fama and French, 2015).

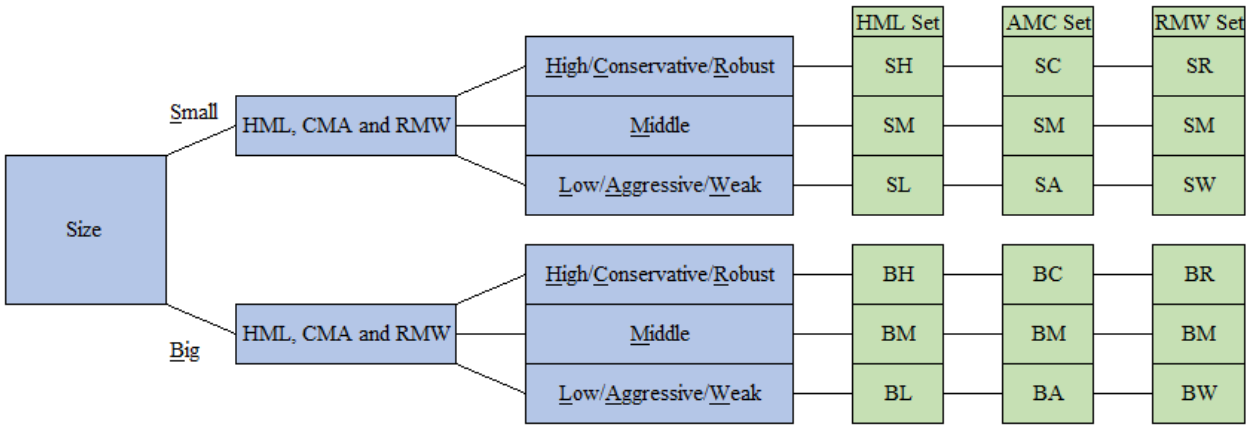


*Figure 3: The figure describes the 2*3 sorted Size-HML, Size-CMA and Size-RMW risk mimicking portfolios, which subsequently were used to construct the explanatory factors.*

SMB_{BTM}, SMB_{CMA} and SMB_{RMW} were calculated for each set of risk mimicking portfolios using the VWR of the portfolios with small corporations minus the VWR of the portfolios consisting of big corporations. Subsequently, an equal-weighted average was calculated (Eq. 7.4) of SMB_{BTM}, SMB_{CMA} and SMB_{RMW} to obtain the SMB factor. (Fama and French, 2015)

$$SMB_{HML} = \frac{SH + SM + SL}{3} + \frac{BH + BM + BL}{3} \quad (7.1)$$

$$SMB_{AMC} = \frac{SA + SM + SC}{3} + \frac{BA + BM + BC}{3} \quad (7.2)$$

$$SMB_{RMW} = \frac{SR + SM + SW}{3} + \frac{BR + BM + BW}{3} \quad (7.3)$$

$$SMB = \frac{SMB_{HML} + SMB_{AMC} + SMB_{RMW}}{3} \quad (7.4)$$

Equation 7.1-4: The Equations show how the SMB risk factor (eq. 7.4) was constructed based on the SMB risk factors of the HML-, CMA- and RMW- risk mimicking portfolio sets (7.1-7.4).

By using each set of portfolios which is presented in figure 3, the return of the remaining explanatory factors can be calculated. The HML factor was calculated by taking the VWR of portfolios with a high book-to-market minus VWR of portfolios with a low book-to-market (Fama and French, 1993). The CMA factor was calculated by taking the VWR of portfolios consisting of stocks classified as conservative minus the VWR of portfolios with stocks classified as aggressive. Finally, the RMW factor was calculated by taking the VWR of portfolios including robust stocks minus the VWR of portfolios including weak stocks (Fama and French, 2015).

$$HML = \frac{SH + BH}{2} - \frac{SL + BL}{2} \quad (8.1)$$

$$CMA = \frac{SC + BC}{2} - \frac{SA + BA}{2} \quad (8.2)$$

$$RMW = \frac{SR + BR}{2} - \frac{SW + BW}{2} \quad (8.3)$$

Equation 8.1-3: The Equations show how each risk mimicking portfolio set from figure 3 was used to construct the HML, CMA and RMW explanatory variables.

4.4.2 Dependent Variables

The dependent variable has been constructed using a similar methodology as in the construction of the explanatory variables. The corporations were independently sorted with respect to the size median and into three groups with regards to the ESG-, Env-, Soc- and Gov- scores separately. The intersection including corporations with same characteristics was put into a portfolio resulting in 6 portfolios for the ESG-score as well as 6 portfolios for each of the individual pillars, hence 24 (6*4) portfolios in total. These portfolios were used as the dependent variables, which is the main

difference compared to the construction of the explanatory variables, where the risk mimicking portfolios were used to calculate the different explanatory variables (Fama and French, 2015).

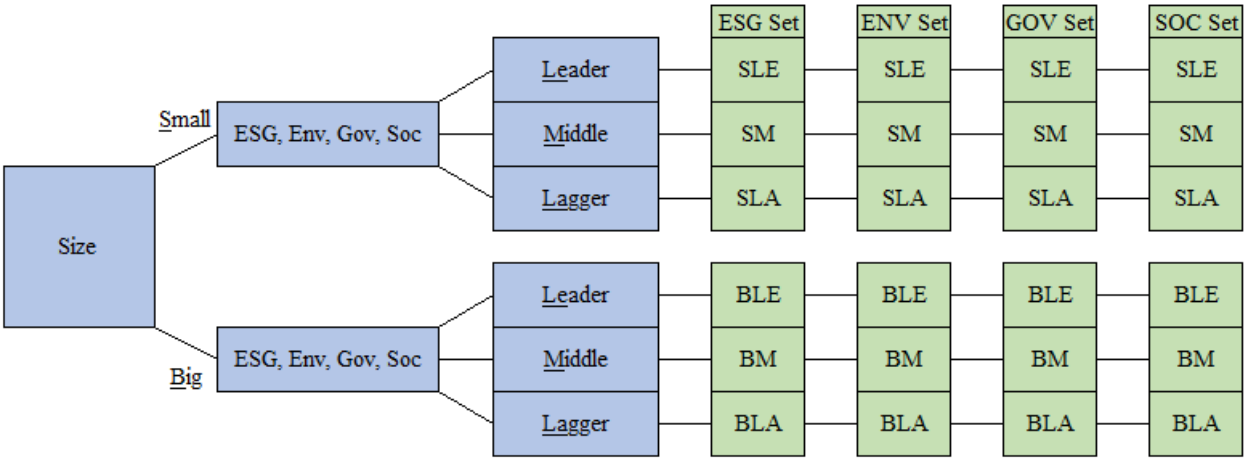


*Figure 4: The figure describes the sorting procedure in the construction of the dependent risk mimicking portfolios based on 2*3 sorted Size-ESG, Size-ENV, Size-Gov and Size-Soc portfolios*

All the portfolios in figure 3 and 4 have been rearranged annually to create updated risk mimicking portfolios. To construct the portfolios both accounting- and market data has been applied. The portfolios were constructed at the end of June year t to ensure that the accounting data for fiscal year t-1 was available.

4.4.3 Essential consideration regarding the sorting Procedure

In line with the Fama-French methodology, the portfolios were constructed by using a 2*3 sorting procedure as described above. FF5FM (Fama and French, 2015) informed that the procedure as well as the chosen breakpoints were arbitrarily selected. Therefore, FF5FM compared the 2*3 sorting procedure to a 2*2 and 2*2*2*2 sorting procedure, where for instance the 2*2*2*2 sorting procedure was based on the intersection between size, B/M, OP (operating profitability) and IR (investment rate). The findings suggested that the significance level of the explanatory factors was unaltered between the different sorting procedures. Therefore, the 2*3 sorting procedure was only applied in this thesis. To construct the dependent portfolios FF5FM used a 5*5 sorting procedure whereas this thesis applies a 2*3 sorting procedure as FF5FM used a larger sample consisting of all stocks at NYSE-, AMEX-, and NASDAQ- indexes.

4.5 Portfolio Evaluation and Model Comparison Metrics

4.5.1 Sharpe Ratio

The Sharpe ratio is one of the most applied metrics to compare the success of investment strategies. The Sharpe ratio takes both the (excess) return and risk into account, as a rational investor would like to receive higher returns for undertaking additional risk. Risk is commonly proxied by volatility within finance, and hence is the Sharpe ratio formula defined as:

$$Sharpe\ Ratio = \frac{Excess\ Return}{SD\ of\ the\ excess\ return} \tag{9}$$

The yearly average Sharpe ratio has been used in this thesis. The yearly Sharpe ratio was calculated by taking the average monthly Sharpe ratio for a given year multiplied by the square root of 12 as the standard deviation rises in direct proportion to the square root of a unit of time i.e slower than the return which increases in direct proportion to time (Bodie, Kane and Maccus, 2014).

4.5.2 Information Criteria

Information criteria have been used to evaluate the model specification. The information criteria capture two competing effects as more explanatory variables are added. That means, the information criteria reward models where the decrease in the residual sum of squares (RSS) outweigh the penalty term, which reflects the loss of degrees of freedom in a complexer model. The information criteria that have been used in this thesis are Akaike information criterion (AIC) and Bayesian information criterion (BIC). BIC gives a higher penalty than AIC, and is hence normally indicating the use of fewer explanatory variables. (Brooks, 2019). AIC and BIC are defined as:

$$AIC = -\ln(\Sigma e^2) + \frac{2k}{N} \tag{10.1}$$

$$BIC = -\ln(\Sigma e^2) + \frac{k}{N} * \ln N \tag{10.2}$$

Where Σe^2 is the sum of squares residuals (residual variance), k is the number of parameters and N is the sample size (Brooks, 2019). A better model corresponds to a lower value on AIC and BIC, which can be concluded from the expression in equation 10.

4.6 Panel Data Regressions

A panel data regression analysis has been conducted in order to establish the relationship between shareholder return and ESG-score, which also enabled a comparison to the ESG- screening results in the portfolio approach. In a panel data setting, one follows a given set of individuals over a specific time frame and its consequently a combination of both cross-sectional data and time-series data. One could compare to the time series regressions in the portfolio approach, where a single portfolio, at a time, was regressed over a given time period. The most important advantage of using a panel data approach is that the data is more effectively organized, which allows one to examine complexer research issues. The effectiveness of the model is reflected in a higher power i.e. more degrees of freedom. A common issue within regression analysis is the omitted variable bias, which could cause biased results and higher variance. An omitted variable bias emerges if an unobserved variable is correlated with an explanatory variable, *endogeneity*, and at the same time has explanatory power on the dependent variable. The fixed effect panel data estimator is able to alleviate the omitted variable bias as the estimator controls for within variation, i.e. time variation within each individual. The random effect panel data estimator controls for the cross-sectional variation between individuals at a given point in time, *time-invariant.* The choice between fixed- and random effect is commonly based on the Hausman test, in which the null hypothesis is corresponding to the random effect. The variables that were used in the panel data analysis are ESG, Env, Soc, Gov, excess return of the market (RMRF), book-to-market (BTM), the logarithm of market capitalisation (LnMC), the logarithm of total assets (LnTA) and profitability margin (PM). The logarithm of total assets and market capitalisation has been taken to assure stationarity. All of the variables are based on market anomalies identified by Fama-French (1993) and Fama-French (2015). The fixed-effect model specification that has been applied in this thesis is:

$$R_{it} = \alpha_i + \mu_i + \beta_i(ESG_{it}) + \beta_i(ENV_{it}) + \beta_i(SOC_{it}) + \beta_i(GOV_{it})$$
$$+\beta_i(RMRF_{it})+\beta_i(BTM_{it})+\beta_i(LMC_{it})+\beta_i(LTA_{it})+\beta_i(PM_{it}) + \sum_{n=2}^{16} YEAR_n + \varepsilon_{it} \quad (11)$$

Where R_{it} is the return for corporation *i* at time *t*, μ_i is the fixed effect related to corporation *i*, the control variables are defined as presented above and $\sum YEAR_n$ is the sum of the year-specific effects.

Inspired by Franzen (2019), three more specifications have been used in order to find out how the model specifications affect the results. The ESG-scores and its pillars have been taken from the same period as in equation 11, lagged one period, this period subtracted by the previous period, *first difference*, and the lagged first difference. The effect of the ESG-score quantity, as well as the change in the ESG-score, will thereby be tested, and also how fast the change is incorporated in the price. In addition, robustness tests before and after the financial crisis have been conducted, which enabled a comparison to the corresponding regressions in the portfolio approach.

4.7 Data Tests

To make sure the ordinary least square (OLS) estimator was the Best Linear Unbiased Estimator (BLUE) and provided reliable test-statistics in the portfolio approach, conventional data tests were conducted. Test for heteroscedasticity was conducted by applying the Breusch Pagan test and White test. To test for autocorrelation, in the residuals, the Durbin-Watson test and Breusch Godfrey test were applied. A skewness and kurtosis test as well as the Shapiro-Wilk test were conducted to ensure a normal distribution in the residuals. Asymptotically, the returns should be normally distributed around the population mean due to the Central Limit Theorem (CLT) (Brooks, 2019). As the data set of this thesis was relatively large, the return data should approach normality due to CLT. An implicit assumption in the application of the OLS estimator is that the explanatory variables are orthogonal to one another, which has been tested by applying a correlation matrix and observing the variance inflation factor (VIF). A too high correlation between explanatory variables generates biased coefficients as the regressors impact one another. There is no clear upper limit but the rule of thumb is 0.8 related to the pairwise correlation matrix and 5 related to the VIF. Finally, the Augmented Dickey-Fuller test and Phillips-Perron test have been conducted on all the variables, but it would be unlikely to find a unit root as returns are not expected to have any particular trend (Brooks, 2014).

In the panel data fixed approach the following data tests were conducted to understand the data characteristics, a groupwise heteroskedasticity test, Wooldridge test for autocorrelation/serial correlation, Pesaran test for weakly cross-sectional dependence i.e. if there is a significant

correlation between the residuals of the corporations. Phillips-Perron test for stationarity and multicollinearity test through the pairwise variable correlation.

4.8 Validity and Reliability

4.8.1 Validity

Validity is divided into *internal validity* and *external validity*. Internal validity concerns how well the methodology of the study measures what the study is intended to measure while external validity points to what extent the results can be generalized in other contexts e.g. at another market (Lundahl and Skärvad, 2009). The authors suggest that the study has high internal validity as the quantitative method design has been applied in a large body of previous academic research. However, the external validity could be questioned as to the STOXX600 and STOXX1800 not necessarily constitute a good proxy of the European market and the market portfolio respectively. Perhaps a more extensive selection of European and global stocks could have been used that reflects the characteristics of these markets. The fact that previous research has provided contradictory evidence dependent upon market and time period is also indicating that the results are hard to generalize.

4.8.2 Reliability

Reliability is related to the existence of unsystematic errors, which could reduce the trustworthiness of the study (Lundahl and Skärvad, 2009). Sustainability is by nature a qualitative phenomenon and the quantification of ESG-scores may not efficiently reflect the sustainability qualities of a corporation. Also, as addressed in legitimacy theory, poor sustainability practices are sometimes uncovered by news organisations. The ESG-score does not address these controversies, as ESG combined score does, which could make the score unreliable. However, the sustainability score is potentially the best quantification of sustainability practices, thus being a good proxy for this study. Another issue that could reduce the reliability is the human failures in the construction of factor mimicking portfolios and filtration of stocks. To cope with this problem, the authors have taken precautionary measures by checking the excel files for faults twice and used excel macros. This should make the study more reliable, but cannot guarantee 100% certainty that the results and

the models are accurate. With regards to reliability, the exclusion of securities for the creation of factor mimicking portfolios is significant. During the first five years of the study period, exclusions were notable due to missing ESG data. As many as 100-150 companies, out of 600, in the STOXX600 were excluded. This exclusion problem disappeared in the following years. In the portfolio approach, one could have used different model specifications as in the panel data approach, for example by lagging the control variables.

5. Result and Analysis

5.1 Descriptive Statistics and Data Tests

To get an overview of the data characteristics, some of the most important data attributes will be presented. Statistical data tests will also be implemented in order to make sure that the data provide reliable results. In table 1, the average monthly excess returns of the sustainability sorted portfolios are presented.

Table 1

*The average compounded monthly excess returns for 6 portfolios 2*3 sorted on Size-ESG, Size-ENV, Size-SOC and Size-GOV. The regressions are based on 192 observations from July 2002 through June 2018.*

%	Lagging	Mid	Leading
Panel A: Size-ESG Portfolios			
Small	0.50	0.69	0.29
Big	0.32	0.32	0.44
Panel B: Size-ENV Portfolios			
Small	0.52	0.66	0.42
Big	0.40	0.34	0.40
Panel C: Size-SOC Portfolios			
Small	0.40	0.78	0.36
Big	0.27	0.33	0.40
Panel D: Size-GOV Portfolios			
Small	0.57	0.57	0.51
Big	0.33	0.11	0.45

Item 1 in the appendix shows the number of financial corporations in the STOXX600 at the end of June on a yearly basis. As seen in the table, the number of financial corporations varies from approximately 120 to 145. The number of financial corporations started to decrease in the financial crisis from 2007 until 2012, and from 2012 the number of financial corporations began to increase again. Item 2 in the appendix shows the number of corporations used in the construction of the explanatory variables. There is a clear upward trend for the entire period, the lowest number in 2002 is around 400 and the lowest number in 2018 is clearly above 450 corporations. The number

of missing corporations from STOXX 600 exceeding the financial corporations are due to missing data. It implies that the number of corporations missing due to lack of data has decreased significantly. In figure 5, the number of corporations applied in the construction of the dependent portfolios is presented. Based on the number of financial corporations in item 1, it is apparent that approximately 100 corporations are missing sustainability data in each dimension, likely to a high extent the same corporations. There is no indication that this gap is significantly decreasing since 2006. If the missing corporations are sharing corporate characteristics the conclusions about the European market would be suffering from a selection bias whereas the result is unbiased if the companies would be missing in a random manner.

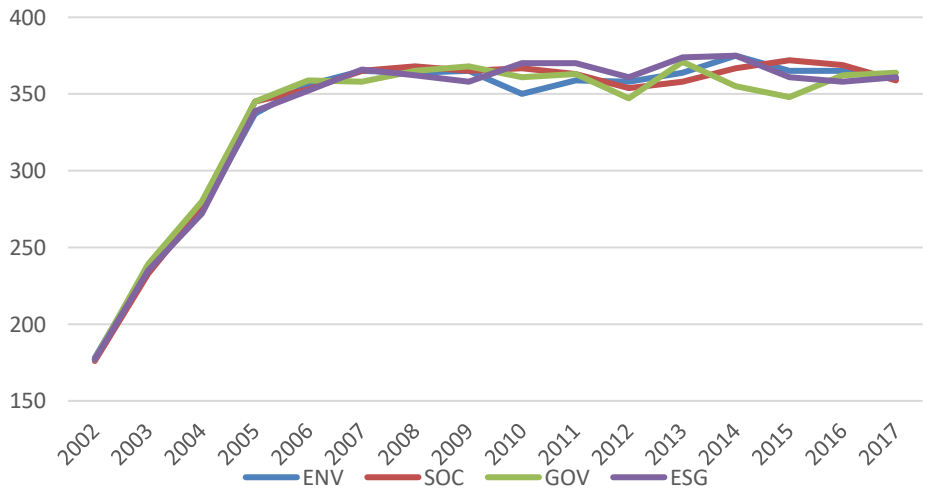


Figure 5: Number of corporations used in the construction of the sustainability portfolios. Missing corporations can be attributed to either exclusion of financial corporations or missing data.

As seen in table 2, the highest mean return is obtained by taking a long position in corporations with a robust (high) earning profile and a short position in a corporation with a weak (low) earnings profile, whereas the lowest mean is related to taking a long position in up momentum corporations and a short position in down momentum corporations. RMW also has a low standard deviation in relation to the mean return. SMB has a higher return than the market portfolio but a significantly higher standard deviation as predicted by CAPM, which is also reflected in the most extreme monthly return observations. The lowest volatility is obtained by investing in accordance with the CMA factor.

Table 2

*The portfolio characteristics of 2*3 sorted portfolios based on Size-ESG, Size-ENV, Size-SOC and ESG-GOV from July 2002 through June 2018 (192 observations). The number of observations (N), monthly mean return (mean), standard deviation (sd), skewness (Sk), Kurtosis (K), minimum monthly return (min) and maximum monthly return (max).*

VARIABLE	**N**	**mean**	**sd**	**Sk**	**K**	**min**	**max**
Explanatory variables							
RMRF	192	0.00438	0.0378	-0.5672	4.1667	-0.121	0.117
SMB	192	0.00557	0.0601	0.7271	10.8782	-0.25	0.373
HML	192	0.00222	0.0224	0.2022	4.6941	-0.0644	0.0978
CMA	192	0.00208	0.018	1.3404	11.6469	-0.0485	0.113
RMW	192	0.00695	0.0178	-0.0027	5.0419	-0.0606	0.0735
UMD	192	0.0017	0.0297	-1.6030	12.8622	-0.199	0.0767

Item 3 in the appendix provides portfolio return characteristics, SLA, SM and BLE structurally receive a higher monthly mean return across the sustainability dimensions. However, the SM portfolio in the social dimension is a conspicuous exception. One can also notice that small corporations tend to receive a higher mean return than big corporations as well as being attached with a higher standard deviation (volatility). SLE tends to be the most skewed portfolio but it's not a problem as the normality assumption of a BLUE OLS estimator is related to the error terms and not normality in data. The descriptive statistics of the panel data fixed effect variables can be seen in item 4 in appendix.

To make sure that the OLS estimator is the Best Linear Unbiased Estimator (BLUE) and provides reliable test-statistics, conventional data tests are conducted. To get an overview of the test structure and the analysis of the data tests, the results of the ESG BLE portfolio within the FF5FM framework are presented in item 5 in the appendix. The Breusch Pagan test is rejecting the null hypothesis of homoscedasticity whereas the White test is not rejecting the null hypothesis of homoscedasticity. In case of contrary evidence, robust standard errors are always applied in order to create reliable test statistics. Both Durbin-Watson and Breusch Godfrey tests are not rejecting the null hypothesis of no autocorrelation. The skewness/kurtosis test indicates no problem with skewness but rejects the null hypothesis of normal kurtosis at 5% significance level. The skewness assumption is the most important as the test-statics values are based on the tails. Furthermore, the Shapiro-Wilk test is not rejecting the null hypothesis of a normal distribution in the residuals. Note

that the correlation matrix and the VIF factor tests for orthogonality of the explanatory variables are generic to all model specifications. The results indicate that the applied set of regressors are not severely correlated. The same conclusion was drawn for the period before and after the financial crisis. As expected in return data, unlike stock price data, there is no trend in the variables as the null hypothesis of stationarity cannot be rejected in either the Augmented Dickey-Fuller test or in Phillips-Perron test. To sum up, the only OLS assumption that was violated was the homoscedasticity assumption and it was solved by using robust standard errors. The other tests behaved in line with a BLUE OLS estimator.

As seen in item 6 in the appendix, all of the five regression specifications, including various sustainability aspects, were highly heteroscedastic and hence robust standard errors were always used. The specifications also portrayed serial correlation and cross-sectional dependence, which may be due to persistent and uncontrolled trends within each corporation and across corporations. To handle this data characteristics various model specifications were used including lags and differences in the sustainability variables as well as a one period lag in the control variables. (Brooks, 2019). The model specifications showed no problem with heteroscedasticity, the highest pairwise variable correlation was between LnMC and LnTA (0.77). In the Phillips-Perron test for stationarity with the null hypothesis of non-stationarity, unit root, the null hypothesis of non-stationarity was always rejected. ESG was closest to being non-stationary with a p-value of 0.9442. In case of violated OLS assumptions, one could also use the feasible generalized least square (FGLS) where weighted sum of squared residuals are minimized. (Brooks, 2019)

5.2 Portfolio Approach

In this thesis, the ESG screening investment strategy is partly evaluated by using multifactor models. It is therefore important to compare the results of a set of multifactor models in order to find the most appropriate model specification. To increase the chance of identifying the best possible model specification, a set of tests and metrics will be applied. By using merely one specification of the multifactor model, the results would just be a comparison to one benchmark investment strategy, for instance, the CAPM benchmark investment strategy is corresponding to investing in a market mimicking portfolio. The validity of the results would consequently depend upon the validity of the model assumptions. In the selection of the optimal model, there is a trade-

off between the increased explanatory power of a complexer model and that the model should be as parsimonious as possible, i.e. adding more factors have to be offset by increased explanatory power. The various model specifications also serve as a robustness check of the relative performance of the 2*3 size-sustainability sorted portfolios.

In table 3, the regression results from July 2002 through June 2018 are presented for various factor combinations, CAPM, FF3FM, C4FM and FF5FM, as well as for the various sustainability dimensions, ESG, ENV, SOC and GOV. By looking at Jensen's alpha and the parameters of the explanatory variables, one is able to conclude that the structure regarding the significance level of the intercepts and the parameters remain similar across the various model specifications and sustainability dimensions. However, the FF5FM deviates in comparison to the other models as almost all Jensen's alphas are significantly different from zero at one percent significance level, insinuating a worse capacity to price the applied portfolios correctly. Jensen's alpha also reveals that the multifactor models, including CAPM, tend to structurally price the SLA, SM and BLE portfolios worse than SLE, BLA and BM. Implying at least a higher ex-post risk-adjusted return by investing in SLA, SM and BLE due to the positive sign on Jensen's alpha. There is no significant difference in return between leading- and lagging corporations as seen in item 7. By observing the significance level of the parameters related to the factors of the FF5FM, it becomes evident that the factors with the highest explanatory power are RMRF, RMW and SMB, while HML and CMA explain the excess return of the applied portfolios worse. In item 8 panel B, it is also seen that all momentum parameters are insignificant. All the factor parameters are frequently significant for leading sustainability corporations independent whether it is a small or big corporation. The factor parameters are also highly significant in, for instance, the BLA portfolio in the ESG- and SOC. sustainability dimensions. Nevertheless, the models are not able to price the dependent portfolios correctly as seen at the alphas in table 3 panel D.

Table 3

*Regressions for six 2*3 sorted Size-ESG, Size-Env, Size-Soc and Size-Gov portfolios, as presented in table 1, based on four different model specifications: CAPM, FF3FM, C4FM and FF5FM. The regressions are based on 192 observations from July 2002 through June 2018. All the factor coefficients are related to the FF5FM as mentioned in Panel D. FF5FM is the model with most factors and is defined as:* $R_{pt} - R_{ft} = \alpha_{pt} + \beta_{pt}(R_{mt} - R_{ft}) + s_{pt}SMB_t + h_{ptHML_t} + r_{pt}RMW_t + c_{pt}CMA_t + e_{pt}$

	ESG			*ENV*			*SOC*			*GOV*		
	Lagging	*Mid*	*Leading*	*Lagging*	*Mid*	*Leading*	*Lagging*	*Mid*	*Leading*	*Lagging*	*Mid*	*Leading*
	Panel A: CAPM intercepts: Rm-Rf											
	α			α			α			α		
Small	0.0038*	0.0053**	0.0024	0.0042*	0.005**	0.0032	0.003	0.006***	0.003	0.0044**	0.0043**	0.0038
Big	0.0025	0.0025	0.0036**	0.0031*	0.0028	0.0032**	0.0023	0.0027	0.0034**	0.0023	0.0015	0.0036**
	Panel B: Three factor model intercepts: Rm-Rf, SMB and HML											
	α			α			α			α		
Small	0.0029*	0.0039**	0.0009	0.0033**	0.0037**	0.0016	0.0021	0.0048***	0.0015	0.0034**	0.0031*	0.0025
Big	0.0022	0.0023	0.0036**	0.0028	0.0026	0.0031*	0.0022	0.0023	0.0033**	0.0018	0.0012	0.0035**
	Panel C: Four factor model intercepts: Rm-Rf, SMB, HML and UMD											
	α			α			α			α		
Small	0.0028*	0.0042**	0.0009	0.003*	0.004**	0.002	0.002	0.0051***	0.0017	0.0033**	0.0031**	0.003
Big	0.002	0.0022	0.0038**	0.0028	0.0027	0.0032**	0.0022	0.0022	0.0035**	0.0018	0.0012	0.0037**
	Panel D: Five factor model coefficients: Rm-Rf, SMB, HML, RMW and CMA											
	α			α			α			α		
Small	0.0065***	0.0068***	0.0054***	0.0068***	0.0065***	0.006***	0.0054***	0.0079***	0.006***	0.0062***	0.0059***	0.0076***
Big	0.0057***	0.0040**	0.0066***	0.0041**	0.0058***	0.0059***	0.0055***	0.0036**	0.0068***	0.0045**	0.0029*	0.0068***
	β			β			β			β		
Small	0.809***	0.777***	0.824***	0.802***	0.811***	0.866***	0.823***	0.781***	0.754***	0.803***	0.812***	0.809***
Big	0.819***	0.850***	0.776***	0.796***	0.802***	0.802***	0.853***	0.837***	0.759***	0.866***	0.802***	0.801***
	s			*s*			*s*			*s*		
Small	0.308***	0.344***	0.276***	0.334***	0.313***	0.148***	0.323***	0.326***	0.264***	0.306***	0.314***	0.339***
Big	0.232***	0.0626	-0.067***	0.0379	-0.0571**	-0.0571**	0.192***	-0.0016	-0.0515*	0.154***	-0.0091	-0.056**
	h			*h*			*h*			*h*		
Small	-0.053	0.114	0.391***	-0.0577	0.104	0.416***	-0.0451	0.111	0.343***	0.0016	0.12	0.107
Big	-0.224*	0.0312	0.206**	-0.132	0.127	0.192**	-0.266**	0.276***	0.143	0.0452	0.154	0.193**
	r			*r*			*r*			*r*		
Small	-0.399***	-0.357***	-0.525***	-0.413***	-0.326***	-0.491***	-0.366***	-0.370***	-0.533***	-0.326***	-0.326***	-0.589***
Big	-0.396***	-0.198*	-0.331***	-0.15	-0.331***	-0.329***	-0.370***	-0.117	-0.402***	-0.295***	-0.192**	-0.362***
	c			*c*			*c*			*c*		
Small	-0.256**	-0.0497	-0.302**	-0.191	-0.142	-0.356**	-0.244**	-0.152	-0.183	-0.167	-0.151	-0.280**
Big	-0.260*	-0.135	-0.289**	-0.0653	-0.389***	-0.158	-0.252*	-0.238**	-0.205*	-0.219	-0.0895	-0.268**

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

In table 4, regressions are run based on exclusively the explanatory factors using each factor as the dependent variable and the remaining as explanatory variables, which gives for instance five regressions based on the FF5FM. An intercept significantly different from zero is indicating that the dependent variable is contributing with explanatory power to the given set of regressors (Fama and French, 2015). The intercepts related to the dependent variables turn out to be overall insignificant in FF5FM, C4FM and FF3FM, although both RMRF and RMW are significant in the FF5FM specification. By testing all possible pairwise combinations of variables, the indication in table 3 was confirmed namely that RMRF and RMW are the only two factors that add explanatory power to one another. A regression with the factors RMRF and RMW on the applied portfolios, however frequently generated significantly positive Jensen's alphas, which is indicating a worse ability to price the applied portfolios than the CAPM. It is important to emphasize that the factor regressions identify the best specified general model, which is best suitable to capture anomalies of securities related to the entire investment universe. The pricing accuracy of different portfolios doesn't necessarily have to be aligned with the best specified general model (Barillas and Shanken, 2015).

Table 4

Factor regressions based on FF5FM, C4FM and FF3FM. Note that the factors are the same in all sustainability dimensions, ESG, Env, Soc and Gov. Each factor is explained by the remaining set of regressors. The factor regressions are based on 192 observations from July 2002 through June 2018.

Factor Regressions

Panel A: FF5FM

VARIABLE	RMRF	SMB	HML	RMW	CMA
RMRF		0.358***	0.0623*	-0.107***	-0.0312
SMB	0.139***		0.058***	-0.0224	-0.0508***
HML	0.310*	0.745***		-0.277***	0.582***
RMW	-0.491***	-0.266	-0.256***		0.187***
CMA	-0.228	-0.958***	0.855***	0.298***	
Constant	0.00681**	0.00618	0.00163	0.00754***	-0.000096
P-Value	(-0.0141)	(-0.168)	(-0.193)	(-0.00)	(-0.926)
R2-Adj	0.169	0.15	0.521	0.497	0.179

Panel B: C4FM

VARIABLE	RMRF	SMB	HML	UMD
RMRF		0.472***	0.105**	0.072
SMB	0.203***		0.00424	-0.161***
HML	0.291**	0.0274		-0.315***
UMD	0.122	-0.634***	-0.192***	
Constant	0.0024	0.00452	0.00207	0.00298
P-Value	(-0.358)	(-0.256)	(-0.186)	(-0.136)
R2-Adj	0.116	0.187	0.091	0.154

Panel C: FF3FM

VARIABLE	RMRF	SMB	HML
RMRF		0.474***	0.0965**
SMB	0.185***		0.0375
HML	0.254**	0.253	
Constant	0.00279	0.00292	0.00159
P-Value	(-0.283)	(-0.483)	(-0.321)
R2-Adj	0.113	0.1	0.037

**** p<0.01, ** p<0.05, * p<0.1*

In table 5, some previous findings are further clarified and additional findings with regard to the performance of various multifactor models are presented. In column 1, the average absolute value of the intercepts, A|ai|, is presented with the corresponding test-statistic and p-value in column 2 and 3 respectively. The intercepts of CAPM and FF3FM are not significantly different from zero, whereas C4FM and FF5FM are significantly different from zero at 5 and 1 percent significance level respectively. A|ai|/A|ri| is defined as the average absolute value of the intercepts divided by the average absolute value of portfolio deviation, which is the sum of the average value of the time series return of portfolio i, where i goes from 1 to the number of included firms N, subtracted by the cross-sectional average of the remaining time series portfolio returns. The ratio is an indication of how much of the dispersion in the return that cannot be captured by the model specification and thereby ends up as Jensen's alpha. For all sustainability dimensions, this ratio ranges from 21,1% to 52%, the worst model is FF5FM followed by CAPM, C4FM and FF3FM. The dispersion with regard to the return of a given portfolio is independent of the model specification and hence is the ratio just a way to put the intercepts in perspective in order to reach a better understanding of the result. In column 5-7, 3 additional model specification metrics are presented to enable an

evaluation of the change in the model specification accuracy as additional factors are included. The average of the metrics is presented in table 5 to get a full overview see item 9 in the appendix. All model metrics are constructed to reward a better explanation of the returns and at the same time penalize a complexer model containing additional factors. As mentioned before, a more parsimonious model is always preferred if the additional factors are not adding enough explanatory power (Bodie, Kane and Marcus, 2014). Due to the construction of the metrics, R2-adjusted penalizes the least, followed by AIC and BIC. All the models favour the FF5FM, followed by FF3FM, C4FM and CAPM. The superior penalty of the BIC metric is, for instance, reflected in the smaller distance between CAPM and FF5FM.

Table 5

A presentation of metrics for various model specifications, CAPM, FF3FM, C4FM and FF5FM, and sustainability dimensions, ESG, Soc, Env and Gov. The model specification metrics are based on 192 observations from July 2002 through June 2018.

	ESG						
	A|ai|	GRS test Stat.	GRS P-value	A|ai|/A|ri|	AIC	BIC	R2-Adj
CAPM	0.0034	4.035	0.072	0.285	-831.0	-824.5	0.64
FF3FM	0.0026	2.172	0.096	0.222	-891.7	-878.7	0.74
C4FM	0.0027	1.826	0.048	0.226	-890.9	-874.6	0.74
FF5FM	0.0058	1.967	0.001	0.496	-908.3	-888.7	0.76
	Social						
	A|ai|	GRS test Stat.	GRS P-value	A|ai|/A|ri|	AIC	BIC	R2-Adj
CAPM	0.0034	5.385	0.014	0.302	-838.7	-832.2	0.64
FF3FM	0.0027	3.058	0.025	0.240	-898.1	-885.0	0.73
C4FM	0.0028	2.471	0.007	0.246	-897.1	-880.8	0.73
FF5FM	0.0059	2.750	0.000	0.520	-914.5	-894.9	0.76
	Environment						
	A|ai|	GRS test Stat.	GRS P-value	A|ai|/A|ri|	AIC	BIC	R2-Adj
CAPM	0.0036	3.394	0.163	0.287	-844.0	-837.5	0.65
FF3FM	0.0028	1.564	0.143	0.228	-901.0	-888.0	0.74
C4FM	0.0030	1.623	0.160	0.237	-900.6	-884.3	0.74
FF5FM	0.0058	1.554	0.003	0.469	-916.9	-897.3	0.76

	Governance						
	A|ai|	GRS test Stat.	GRS P-value	A|ai|/A|ri|	AIC	BIC	R2-Adj
CAPM	0.0033	4.805	0.136	0.271	-850.7	-844.1	0.66
FF3FM	0.0026	1.609	0.154	0.211	-907.2	-894.1	0.75
C4FM	0.0027	1.584	0.147	0.219	-906.0	-889.7	0.75
FF5FM	0.0056	1.648	0.000	0.461	-922.6	-903.1	0.77

It is not apparent what model to select as the Jensen's alpha test in table 3, factor regression test in table 4 and the model specification evaluation in table 5 are partly providing contradictory evidence. First of all, one could exclude the C4FM as the model specification metrics are not indicating additional explanatory power in comparison to the FF3FM and the more parsimonious FF3FM is performing similarly with regard to Jensen's alpha and the factor regressions. The highly significant mispricing of the applied portfolios by the FF5FM makes the model unsuitable, even though the information criteria, AIC and BIC, favour FF5FM. The lowest average value of the absolute intercepts is related to the FF3FM and the information criteria favoured FF3FM significantly over CAPM. It can be seen that the difference between CAPM and FF3FM is much bigger than the difference between FF3FM and FF5FM. The FF3FM also has lower abnormal return than the CAPM measured by Jensen's alpha. On the other hand, the FF3FM factor regressions in table 4 don't generate significant intercepts which would be in favour of the more parsimonious CAPM. It is thereby clear that the factor regressions in table 4 favour a more parsimonious model than the model specification tests, GRS, R2-adjusted, AIC and BIC, in table 5. The CAPM and FF3FM are providing very similar results, as seen in table 3, which means that the conclusion based on Jensen's alpha is more reliable.

The ex-post Sharpe ratios related to the applied portfolios are analysed to provide complementing evidence to Jensen's alphas in order to establish how different ESG screening strategies have performed in comparison to one another. As seen in table 6, an investor wouldn't prefer to invest in exclusively leading- or lagging sustainability corporations in the pursuit of a higher risk-adjusted return. The best strategy is to invest in relatively small corporations with a middle ESG-score, the strategy is most auspicious in the social dimension. The second-best performing investment strategy is to invest in a BLE sustainability portfolio. As seen in table 3, smaller corporations tend

to have a higher monthly mean return but BLE is receiving a higher risk-adjusted return than both SLE and SLA due to lower volatility. One is not able to draw the conclusion that it is always better to invest in leading- or lagging sustainability corporations, as the BLE portfolios are outperforming the BLA portfolios whereas SLA portfolios are outperforming SLE portfolios. The performance-based on the ex-post Sharpe ratio favours the same portfolios, SLA, SM and BLE, which frequently generated a significantly higher risk-adjusted return, Jensen's alpha, at 5 percent significance level. Note that the relative performance of the portfolios remains the same between different sustainability dimensions as seen in table 6. The information criteria also favour the portfolios with the highest ex-post risk-adjusted return with the exception of BM, see item 9 in the appendix.

Table 6

*Sharpe ratio for six 2*3 sorted Size-ESG, Size-Env, Size-Soc and Size-Gov portfolios. Note that the ex-post portfolio Sharpe ratio is independent of the model specification. The Sharpe ratios are based on 192 observations from July 2002 through June 2018.*

	BIG			**SMALL**			
	LE	MID	LA	LE	MID	LA	Average
ESG	0.656	0.517	0.505	0.465	0.690	0.599	0.572
ENV	0.616	0.572	0.565	0.507	0.687	0.626	0.595
SOC	0.628	0.557	0.474	0.501	0.775	0.537	0.578
GOV	0.646	0.459	0.529	0.559	0.620	0.651	0.578
Average	0.636	0.526	0.518	0.508	0.693	0.603	

5.3 Panel Data Approach

As seen at the panel data fixed effect regressions in table 7, there is a significantly negative relationship between the stock return and the sustainability score. The negative relationship persists independent of the number of lags or if the change of the sustainability score is used. However, the significance level of the sustainability pillars from period t and the first difference (Δt) are higher than the significance level of the sustainability pillars from period t-1 and the lagged first difference ($\Delta t\text{-}1$). It means that the sustainability information is affecting the shareholder returns in both period t and in period t-1, but that the impact is more significant closure to the sustainability score disclosure. In a fully (semi-) efficient market, the whole sustainability effect would have been incorporated in the price directly. The governance score is insignificant in both

period t-1 and in the lagged first difference ($\Delta t\text{-}1$) regressions and is thereby the pillar that is incorporated in the stock price the fastest. The ESG-score is significantly negative in all specifications. The absolute value is higher in period t-1 and in the lagged first difference ($\Delta t\text{-}1$) regressions, even though the standard deviation is higher, which implies a decreased significance level. It is not possible to identify exactly which pillars that drive the shareholder returns as the significance level of the pillars varies across the model specifications. The correlation between the pillars, seen in item 6 in the appendix, could be the explanation why the parameter of a single regression pillar (column 3-5) is more significant than the parameters of the pillars as all are used as regressors (column 2). This occurs in Table 7A, 7C and 7D. The panel data results are indicating an overall negative return by investing in leading sustainability corporations, whereas the portfolio approach indicated no significant difference, see item 7 in the appendix.

Table 7

Panel data fixed effect regressions are based on the period from July 2002 through July 2018 with four different model specifications. In Table A is the sustainability scores from period t, in table B is the sustainability scores from t-1, in Table C is the sustainability scores constructed by taking the first difference and Table D is the sustainability scores constructed by taking the lagged first difference. As seen in column 1-5 in table A-D, different combinations of the sustainability scores are used. The relationship between shareholders return and the parameters related to ESG and its pillars are just presented in table 7. A set of control variables are used based on the market anomalies found by Fama-French, see item 13 in the appendix for the comprehensive results.

Table A					
Variable	1A	2A	3A	4A	5A
ESG_t	-0.013***				
	(0.003)				
ENV_t		-0.004	-0.006***		
		(0.003)	(0.002)		
SOC_t		-0.004		-0.006***	
		(0.003)		(0.002)	
GOV_t		-0.005**			-0.007***
		(0.002)			(0.002)
Obs.	85,361	74,577	85,216	83,972	85,085
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

Table B					
Variable	1C	2C	3C	4C	5C
ΔESG_t	-0.016***				
	(0.003)				
ΔENV_t		-0.010***	-0.012***		
		(0.003)	(0.002)		
ΔSOC_t		-0.003		-0.008***	
		(0.003)		(0.002)	
ΔGOV_t		-0.004*			-0.005***
		(0.002)			(0.002)
Obs.	77,821	70,014	77,628	76,216	77,522
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

Table C					
Variable	1B	2B	3B	4B	5B
ESG_{t-1}	-0.022***				
	(0.008)				
ENV_{t-1}		0.006	-0.002		
		(0.007)	(0.006)		
SOC_{t-1}		-0.023***		-0.016***	
		(0.007)		(0.006)	
GOV_{t-1}		-0.004			-0.005
		(0.006)			(0.006)
Obs,	77,821	70,014	77,628	76,216	77,522
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

Table D					
Variable	1D	2D	3D	4D	5D
ΔESG_{t-1}	-0.036***				
	(0.009)				
ΔENV_{t-1}		-0.027***	-0.030***		
		(0.007)	(0.006)		
ΔSOC_{t-1}		-0.012		-0.023***	
		(0.008)		(0.007)	
ΔGOV_{t-1}		0.000			-0.001
		(0.006)			(0.006)
Obs.	73,650	69,344	73,464	72,288	73,417
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

**** p<0.01, ** p<0.05, * p<0.1, TFE = Time Fixed Effect, FFE = Firm Fixed Effect*

5.4 Discussion of Theories

There is no evidence indicating that the market is weak-efficient in the sense that one is able to obtain an abnormal return by picking stocks based on their previous market value development, which can be observed at the highly insignificant momentum factor, see item 8 in the appendix. The results indicate that the market may not be fully semi-efficient as the RMW factor has high explanatory power, enabling one to define RMW as a market anomaly. In addition, three out of six of the applied portfolios, SLA, SM and BLE, frequently outperform the market benchmark portfolio with regard to Jensen's alpha and ex-post Sharpe ratio, which shouldn't be the case in a fully semi-efficient market. The negative relationship between shareholder return and sustainability score, in the panel data approach, is also indicating that the market is not able to price the publicly available sustainability information accurately. The sustainability score of a corporation could of course influence the total value of equity but a systematic underperformance with regard to shareholder returns of high-sustainability corporations is not compatible with an (semi-) efficient market, unless the corporations are not differing with regard to uncontrolled corporate characteristics which in turn could influence the risk. Nevertheless, the evidence of an inefficient market, in the portfolio approach, is much weaker than Fama-French (1993), Carhart (1997) and Fama-French (2015), although they didn't examine sustainability portfolios. In the portfolio approach, neither a leading- nor lagging sustainability screening strategy was found to be able to generate an abnormal return, which is in line with the semi-efficient market hypothesis. As seen in item 10 and 11, the market is not (semi-) efficient before the financial crisis and highly (semi-) efficient after the financial crisis.

It is hard to fully evaluate the legitimacy- and stakeholder theory due to lack of information about the market participants as well as about the potential outcome (return) if the corporation would have acted differently. Contrary to the legitimacy theory, shareholders don't seem to reward sustainable corporations with a higher return in any structural way, at least not if one assumes that there are no fundamental differences in corporate characteristics in the applied portfolios apart from the applied sorting criteria. In the panel data approach, investors are rather penalized by investing in high-sustainability corporations.

There could be effects that are impacting the result of sustainability undertakings in different directions and hence make the result hard to derive. Corporations with a high sustainability score may be more attractive to investors but have less resources for other internal investments, which otherwise may have impacted the market value positively. Furthermore, one cannot draw any conclusion whether the corporations in general have an ethical- or managerial perspective based on the findings of this thesis. There does not necessarily have to be a trade-off between the ethical- and managerial perspective with regard to sustainability undertakings, for instance in this thesis big companies are able to invest in sustainability and at the same time perform better regarding shareholder returns

5.5 Robustness Tests of Portfolio- and Panel Data Approach

As seen in item 10 in the appendix, the portfolio approach before the financial crisis shows significantly higher Jensen's alphas than during the full period. The model with the least significance is still the CAPM, but the significance level is still overall higher and in particular the SLE portfolio tends to have changed from insignificant to significant. The relatively positive significance level of SLA, SM and BLE in the full period was also prevalent in the sub-period before the financial crisis. As seen in item 11 in the appendix, the significance level of Jensen's alphas has almost disappeared in the period after the financial crisis, independent of the model specification. The SML and RMW factors are also less significant after the financial crisis than before the financial crisis, indicating that these two market anomalies have become less prevalent. The structural shift from an increased significance level before the financial crisis to a decreased significance level after the financial crisis regarding Jensen's alpha could be explained by a more (semi-) efficient market. It may be partly attributed to increased consideration of sustainability-related issues in investment strategy.

As seen in item 12 in the appendix, the panel data fixed effect regression before the financial crisis is still generating a negative shareholder return using a positive ESG screening strategy. However, the significance level is relatively low, for instance, it is only the one month lagged ESG-score that is significant whereas all governance parameters are insignificant. The lagged model is still generating significant parameters of the social and environment pillars at a 5% significance level.

As seen in item 13 in the appendix, the explanatory power of the sustainability variables after the financial crisis is significantly negative in almost all of the specifications. The only exception is the parameters of the sustainability variables related to the lagged first difference specification. The governance factor is once again less significant than the social and environment factors.

To sum up, the positive risk-adjusted return in the portfolio approach has gone from significantly positive before the financial crisis to almost disappear after the financial crisis. The panel data approach shows more significant negative returns after the financial crisis in comparison to before the financial crisis. The decreased risk-adjusted return in portfolio approach could hence be reflected in a more negative return in the panel data approach and vice versa. The positive ESG screening strategy performs at least in both the portfolio and panel data approach worse after the financial crisis than before.

5.6 Comparison to Previous Research

Stevensson (2020) highlighted that the majority of the previous research has not indicated a positive risk-adjusted return by investing in high ESG-score corporations and that rather sin-stocks (lagged sustainability corporations) have historically generated a higher risk-adjusted return. That lagged sustainability corporations outperform leading sustainability corporations are in line with the panel data fixed effect results, whereas the portfolio approach gave an insignificant difference in return between leading- and lagging sustainability corporations. Darren et. al (2012) and Galema et. al. (2008) also found an insignificance difference in return using a leading- or lagging sustainability screening strategy, by applying the portfolio approach on the American market. There is no indication that a positive sustainability screening strategy would generate a higher risk-adjusted return, apart from the positive risk-adjusted return of the BLE portfolio in the portfolio approach. This thesis is therefore providing contrary evidence in comparison to Robert et. al. (2014) who got negative abnormal returns using a positive sustainability screening strategy on the American market. Franzen (2019) examined the difference in return between high- and low ESG corporations using the S&P 500 index. The findings indicated a significantly negative risk-adjusted return by investing in high sustainability corporations, whereas the panel data fixed effect regression was overall insignificant. This thesis also got a discrepancy between the results in the

portfolio approach and panel data approach, even though the result differed. Yea Eun Jang (2019) studied the STOXX600 from 2002 to 2017 using FF3FM and got a negative risk-adjusted return by investing in high ESG corporations in comparison to low ESG corporations. In contrast to Yea Eun Jang (2019), this thesis got an insignificant difference between high- and low sustainability companies independent of the sustainability dimension.

6. Conclusion

In the portfolio approach, the multifactor model comparison indicated that the CAPM or the FF3FM are more suitable to price the applied portfolios than the C4FM and FF5FM. Independent of model specification the SLA, SM and BLE portfolios generated a positive risk-adjusted return measured by Jensen's alpha whereas SLE, BLA and BM generated an insignificant Jensen's alpha. The SLA, SM and BLE were also attached with the highest value on the Sharpe ratio metric. However, there was an insignificant difference in return between high- and low sustainability corporations. The results were overall robust across all of the sustainability dimensions. The panel data fixed effect regressions indicated a significantly negative return by investing in leading sustainability corporations in comparison to lagging sustainability corporations. The panel data fixed effect results were also overall robust independent of the sustainability dimension. In the portfolio approach the market was more (semi-) efficient after the financial crisis as Jensen's alphas got less significant and in the panel data fixed effect regressions the negative relationship between the sustainability score and market value became even more negative after the financial crisis than before. There is no support that an increased ESG-score results in a higher shareholder return through an increased legitimacy. The interest of different stakeholders therefore got to be balanced by the managers.

For future research, it would be interesting to compare markets by using a broader index for example the Stoxx 1800 global index to find out if the performance of an ESG-screening strategy differs across markets. One could also either use another European index or data provider of the sustainability score. In addition, one could analyse how the relationship between ESG-score and shareholder return is affected by the prevalent business cycle as well as try to deeper understand what drives the relationship between ESG-score and shareholder return.

7. References

Barillas, F. and Shanken, J. (2015). Which Alpha? National Bureau of economic research, Nber working paper series. Available online: https://www.nber.org/papers/w21698.pdf [15 May 2020]

Barnett, M. L. and Salomon, R. M. (2011). Does it pay to be really good? Addressing the shape of the relationship between social and financial performance. *Strategic Management Journal*, 33(11), 1304–1320. Available online: https://www.researchgate.net/publication/228248940_Does_it_Pay_to_Be_Really_Good_Addressing_the_Shape_of_the_Relationship_between_Social_and_Financial_Performance [29 April 2020]

Berk, D. and DeMarzo, P (2013). *Corporate Finance.* Pearson Education. Boston.

Bodie, Z., Kane, A. and Marcus, A.J. (2014). *Investments*. McGraw-Hill Education. New York

Borglund, T., De Geer, H. and Sweet, S. (2012). *Corporate Social Responsibility - En guide till företags ansvar*. Stockholm: Sanoma Utbildning

Brooks, C. (2019). *Introductory Econometrics For Finance*. Cambridge University Press, 4th Edition

Bryman, A. and Bell, E. (2017). *Företagsekonomiska forskningsmetoder* . Upplaga 3. Liber. Stockholm.

Byström, H (2010). *Finance - Markets, Instruments & Investments*. Second edition. Studentlitteratur. Lund

Carhart, M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance.* Vol. 52, Issue 1, March 1997. Pages 57-82 Available online: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1997.tb03808.x [12 May 2020]

Cochran, P. and Wood, R. (1984). Corporate Social Responsibility and Financial Performance. *The Academy of Management Journal*, Vol. 27, No. 1 (Mar. 1984), pp. 42-56 Available online: https://www.academia.edu/10764812/Corporate_Social_Responsibility_and_Financial_Performance [29 April 2020]

Council Directive 2014/95/EU, (2014a). Directive 2014/95/EU of the European Parliament and of the Council of 22 October 2014 amending Directive 2013/34/EU as regards disclosure of non-financial and diversity information by certain large undertakings and groups Text with EEA relevance. (2014) Official Journal, p6 - p7. Available online: https://eur-lex.europa.eu/eli/dir/2014/95/oj [17 May 2020]

Council Directive 2014/95/EU, (2014b). Directive 2014/95/EU of the European Parliament and of the Council of 22 October 2014 amending Directive 2013/34/EU as regards disclosure of non-

financial and diversity information by certain large undertakings and groups Text with EEA relevance. (2014) Official Journal, p9. Available online: https://eur-lex.europa.eu/eli/dir/2014/95/oj [17 May 2020]

Darren D. Lee, Robert W. Faff and Saphira A.C. Rekke (2012). Do high and low-ranked sustainability stocks perform differently? University of Queensland Business school. Available online: https://www-emerald-com.ludwig.lub.lu.se/insight/content/doi/10.1108/18347641311312267/full/pdf [14 May 2020]

Deegan, C. and Unerman, J. (2011). *Financial Accounting Theory.* Second European ed. McGraw-Hill.

Elkington, J (1997). *Cannibals with Forks: The Triple Bottom Line of 21st Century Business*, Oxford, Capstone

Edmans, A. (2011). Does the Stock Market Fully Value Intangibles? Employee Satisfaction and Equity Prices. University of Pennsylvania. Available online: https://repository.upenn.edu/cgi/viewcontent.cgi?article=1046&context=fnce_papers [12 April 2020]

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, Vol. 25, No. 2, May 1970, pp. 383-417. Available online: http://efinance.org.cn/cn/fm/Efficient%20Capital%20Markets%20A%20Review%20of%20Theory%20and%20Empirical%20Work.pdf [06 April 2020]

Fama, E. and French, K. (1992). The Cross-Section of Expected Stock Returns. *The Journal of finance*. Vol. 47, Issue2, June 1992, Pages 427-465. Available online: https://doi.org/10.1111/j.1540-6261.1992.tb04398.x [04 April 2020]

Fama, E. and French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*. Vol. 33, 1993, Pages 3-56. Available online: https://rady.ucsd.edu/faculty/directory/valkanov/pub/classes/mfe/docs/fama_french_jfe_1993.pdf [14 May 2020]

Fama, E. and French, K. (2015). A five factor asset pricing model. *The Journal of Financial Economics*, Vol. 116, Issue 1, April 2015, Pages 1-22. Available online: https://doi.org/10.1016/j.jfineco.2014.10.010 [04 April 2020]

Franzén, F. (2019). ESG Portfolios and Stock Returns: An analysis of ESGs effect on financial performance. Master Thesis in Economics, Lund University. Available online: http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=8981910&fileOId=8981922 [15 May 2020]

Friedman, M. (1970). The social responsibility of business is to increase its profits, *New York Times Magazine*, 13 September, 32-33, 122-124. Available online: http://umich.edu/~thecore/doc/Friedman.pdf [04 April 2020]

Freeman, R. (1984). *Strategic management: A stakeholder perspective*. Boston: Pitman.

Galema, R., Plantinga, A. and Scholtens, B. (2008). The stocks at stake: Return and risk in socially responsible investment. *Journal of Banking & Finance*. Vol. 32, Issue 12, December 2008, Pages 2646-2654. Available online: https://www.sciencedirect.com/science/article/abs/pii/S0378426608001325?via%3Dihub [14 May 2020]

Jensen, M. (1967). The Performance of Mutual Funds in the Period 1945-1964. *Journal of Finance*, Vol. 23, No. 2, pp. 389-416. Available online: https://ssrn.com/abstract=244153 or http://dx.doi.org/10.2139/ssrn.244153 [15 April 2020]

Jensen, M and Meckling, W. (1976) Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics*, October, 1976, Vol 3, No. 4, pp. 305-360.
Available online: https://www2.bc.edu/thomas-chemmanur/phdfincorp/MF891%20papers/Jensen%20and%20Meckling%201976.pdf [25 May 2020]

Landi, G. and Sciarelli, M. (2019). Towards a more ethical market: the impact of ESG rating on corporate financial performance. *Social responsibility journal*, Vol 15 No. 1 2019, pp. 11-27. Available online: https://www-emerald-com.ludwig.lub.lu.se/insight/content/doi/10.1108/SRJ-11-2017-0254/full/pdf?title=towards-a-more-ethical-market-the-impact-of-esg-rating-on-corporate-financial-performance [02 April 2020]

Lins, K., Servaes, H. & Tamayo, A. (2017). Social Capital, Trust, and Firm Performance: The Value of Corporate Social Responsibility during the Financial Crisis
The Journal of Finance. Vo.72 No.4
Available online: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12505 [22 May 2020]

Lundahl, U. and Skärvad, P-H. (2009). *Utredningsmetodik för samhällsvetare och ekonomer*, studentlitteratur. Lund

Minutolo, M., Kristjanpoller, W. and Stakeley, J. (2019). Exploring environmental, social, and governance disclosure effects on the S&P 500 financial performance. Business strategy and the Environment. 2019;28: pp. 1083-1095. Available online: https://onlinelibrary-wiley-com.ludwig.lub.lu.se/doi/abs/10.1002/bse.2303 [05 April 2020]

Mitchell, R., Agle, B. and Wood, D. (1997). Toward a Theory of Stakeholder Identification and Salience: Defining the Principle of Who and What Really Counts. *The Academy of Management Review*. Vol. 22, No. 4 (Oct. 1997), pp. 853-886. Available online: https://www.jstor.org/stable/259247?seq=1#metadata_info_tab_contents [17 May 2020]

PRI - Principles of responsible investment (2020), https://www.unpri.org/ [14 May 2020]

Robert G. Eccles, Ioannis Ioannou, George Serafeim (2014). The Impact of Corporate Sustainability on Organizational Processes and Performance. *Management Science* Vol. 60, No. 11. Available online: https://www.hbs.edu/faculty/Publication%20Files/SSRN-id1964011_6791edac-7daa-4603-a220-4a0c6c7a3f7a.pdf [14 May 2020]

Stevenson, David (2020). Are ESG and sustainability the new alpha mantra?, *Financial Times*, 1 April. Available online: https://www.ft.com/content/6cee0b48-7760-46a5-9759-243aaaff7f8a [14 May 2020]

STOXX (2020), https://www.stoxx.com/index-details?symbol=SXXP [15 May 2020]

United Nations (2015). Transforming our world: The 2030 Agenda for Sustainable Development. New York: United Nations. Available online: https://sustainabledevelopment.un.org/post2015/transformingourworld [16 May 2020]

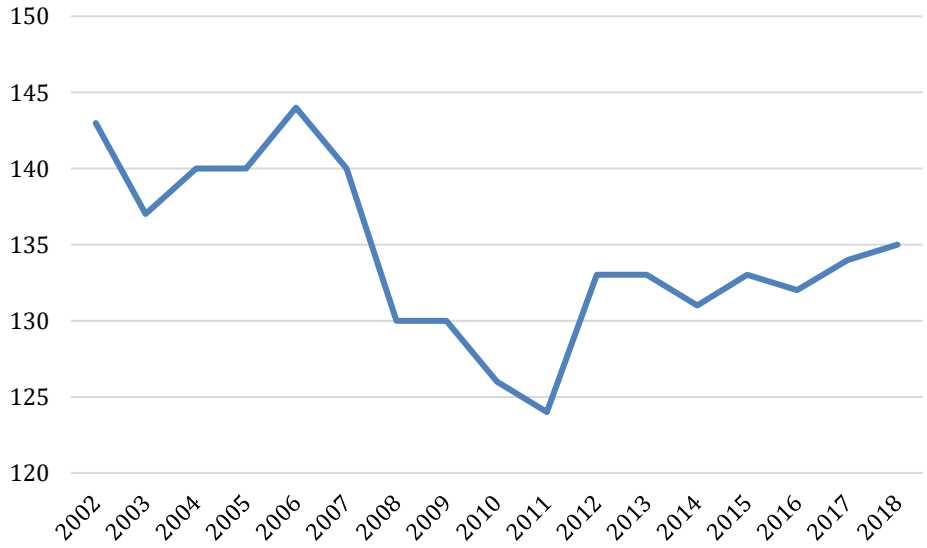
Thomson Reuters (2018). Thomson Reuters ESG-scores,
Available online: http://zeerovery.nl/blogfiles/ESGscores-methodology.pdf [12 April 2020]

Yea Eun Jang (2019). Do ESG scores matter in the market?, Master thesis in finance, Lund University. Available online:
http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=8981480&fileOId=8981488 [14 May 2020]

8. Appendix

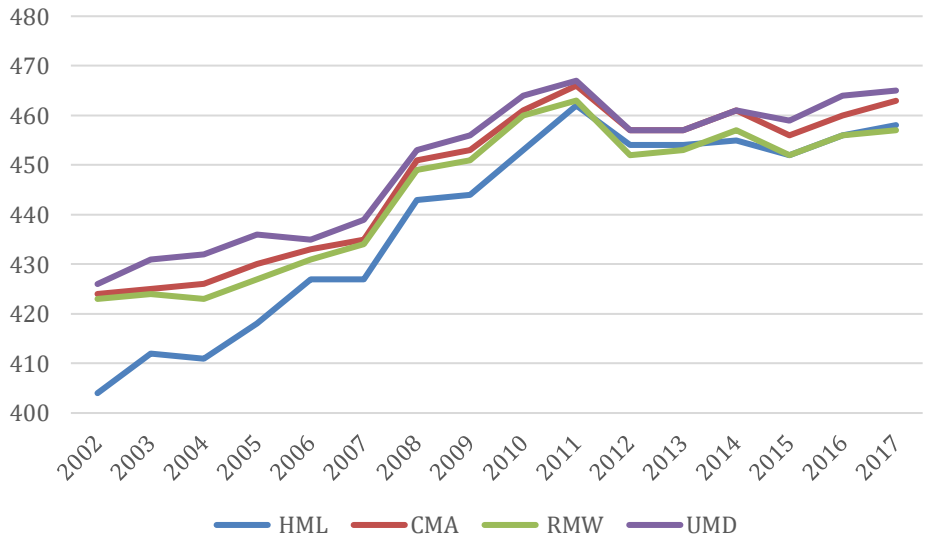
Item 1

The number of financial corporations that hence are excluded from the Stoxx Europe 600 index. The number was observed at the end of June on a yearly basis between 2002 and 2018



Item 2

The number of corporations used in the construction of the independent variables. Missing corporations of the Stoxx Europe 600 can either be attributed to exclusion of financial corporations or missing data. Note that SMB is constructed as an average of HML, CMA and RMW and that RMRF is constructed based on the return data.



Item 3

*The portfolio characteristics of 2*3 sorted portfolios based on Size-ESG, Size-ENV, Size-SOC and ESG-GOV from July 2002 through June 2018. The number of observations (N), monthly mean return (mean), standard deviation (sd), skewness (Sk), Kurtosis (K), minimum monthly return (min) and maximum monthly return (max).*

VARIABLE	N	mean	sd	Sk	K	min	max
ESG							
BLE	192	0.00718	0.0376	-0.7295	4.5694	-0.129	0.114
BM	192	0.00653	0.0419	-0.3576	6.0265	-0.151	0.188
BLA	192	0.00682	0.0477	-0.6938	10.4450	-0.251	0.245
SLE	192	0.00718	0.0532	-0.4529	5.3397	-0.206	0.216
SM	192	0.00978	0.0492	0.2518	10.9617	-0.187	0.308
SLA	192	0.00829	0.0485	-0.7559	7.7707	-0.232	0.22
ENV							
BLE	192	0.00693	0.0385	-0.6291	4.4257	-0.123	0.116
BM	192	0.00665	0.0406	-0.7994	6.2581	-0.164	0.16
BLA	192	0.00724	0.0439	-0.3383	4.8130	-0.155	0.179
SLE	192	0.00803	0.0547	0.3352	10.0519	-0.199	0.332
SM	192	0.00952	0.0488	-0.1003	8.7074	-0.204	0.272
SLA	192	0.00852	0.0471	-0.7684	6.6782	-0.219	0.19
SOC							
BLE	192	0.00695	0.0376	-0.7266	4.8487	-0.131	0.11
BM	192	0.00657	0.0403	-0.5815	4.6276	-0.133	0.14
BLA	192	0.00657	0.0474	-0.7699	10.3793	-0.252	0.237
SLE	192	0.00745	0.0505	-0.3567	4.6435	-0.173	0.199
SM	192	0.0105	0.0482	0.1565	9.9232	-0.182	0.291
SLA	192	0.00761	0.049	-0.7381	8.5908	-0.243	0.234
GOV							
BLE	192	0.00734	0.0388	-0.6655	4.5035	-0.134	0.119
BM	192	0.00516	0.0384	-0.7016	4.6240	-0.13	0.114
BLA	192	0.00671	0.0458	-0.3237	6.7843	-0.173	0.222
SLE	192	0.00856	0.0534	-0.2366	10.3685	-0.26	0.3
SM	192	0.00885	0.0484	-0.3180	7.0114	-0.194	0.24
SLA	192	0.00879	0.0477	-0.3985	8.0327	-0.203	0.248

Item 4

The descriptive statistics of the variables used in the panel data regressions between July 2002 and June 2018 (192 time periods). In the table, the number of observations (N), the arithmetic monthly mean return (mean), the monthly standard deviation (sd), the monthly minimum value (min) and the monthly maximum value (max) are presented. The first row "Return" is the dependent variable in the panel data analysis.

Panel A: Descriptive statistics

VARIABLES	N	mean	sd	Sk	K	min	max
Return	128,382	1.175	10.670	1.243	25.450	-97.930	253.100
ESG	101,738	48.090	21.080	0.018	2.114	0.327	95.130
ENV	101,445	44.180	28.050	-0.071	1.847	0.000	98.730
SOC	100,154	48.490	25.100	0.077	1.936	0.000	98.690
GOV	101,419	50.660	22.410	-0.069	2.070	0.818	99.580
RMRF	171,840	0.300	3.650	-0.655	4.102	-11.411	10.819
LnMC	140,969	21.910	1.498	-0.396	5.044	9.534	26.260
LnTA	159,575	22.050	1.549	-0.110	3.725	14.310	26.770
BTM	140,807	12	1115	94	8814	-381	108417
PM	138,501	0	4	-12	2311	-248	195
Number of Companies	755	755	755	755	755	755	755

Item 5

The data tests are presented to provide an overview of the conducted tests. The portfolio specific tests are related to the ESG BLE portfolio. All regressions are tested for heteroscedasticity, autocorrelation, normality, multicollinearity and stationarity.

Panel A: Heteroscedasticity

	Chi2	P-Value
Breusch Pagan test	7.5	0.0062
White test	19.19	0.5093

Panel B: Autocorrealtion

Durbin Watsson	D-statistics (6, 192)	DL	4-DU	4-DL	4-DU
	2.23	1.51	1.72	2.49	2.28

Breusch Godfrey	Chi2	P-value
	2.74	0.10

Panel C: Normality test

Skewness/Kurtosis test	Skewness P-Value	Kurtosis P-Value	Chi2 Joint	P Value Joint
	0.98	0.05	3.95	0.14

Shapiro-Wilk W test	Z-Statistics	P-Value
	0.23	0.41

Panel D: Multicollinearity

Correlation Matrix	RMRF	SMB	HML	CMA	RMW	UMD
RMRF	1					
SMB	0.32	1				
HML	0.2	0.15	1			
CMA	-0.04	-0.12	0.65	1		
RMW	-0.33	-0.24	-0.21	0.1	1	
UMD	-0.06	-0.33	-0.27	-0.04	0.31	1

VIF	
RMRF	1.23
SMB	1.20
HML	2.13
RMW	2.03
CMA	2.03

Panel E: Stationarity

Augmented Dickey Fuller	BLE	RMRF	SMB	HML	RMW	CMA	UMD
test statistics	-13.85	-12.98	-12.17	-11.48	-12.89	-11.79	-13.69
P-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Phillips-Perron test	BLE	RMRF	SMB	HML	RMW	CMA	UMD
test statistics	-13.85	-12.91	-12.17	-11.54	-12.91	-11.82	-13.66
P-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Item 6

The panel data fixed effect data tests for the period from July 2002 through June 2018 are shown. The data is tested for heteroscedasticity, Wooldridge serial correlation, Pesaran test for weakly cross sectional dependence, multicollinearity and Phillips-Perron test for non-stationarity.

Regression Specification:	**ESG**	**ENV. SOC and GOV**	**ENV**	**SOC**	**GOV**
Panel A: Heteroscedasticity					
Groupwise heteroscedasticity Test	0.000	0.000	0.000	0.000	0.000
Panel B: Serial Correlation					
Wooldridge Test for Serial Correlation	0.000	0.000	0.000	0.000	0.000
Panel C: Crossectional Dependence					
Pesaran Test for Weakly Cross-Sectional Dependence	0.000	0.000	0.000	0.000	0.000
Panel D: Stationarity Test					
Phillips-Perron Unit-Root Test					
ESG	0.9442				
ENV	0.9292				
SOC	0.6083				
GOV	0.000				
RMRF	0.000				
BTM	0.000				
LnMC	0.000				
LnTA	0.000				
PM	0.000				

Panel E: Multicollinearity Test ESG Specification

Correlation	ESG	RMRF	LMC	LTA	BM	PM
ESG	1					
RMRF	0.04	1				
LnMC	0.44	0.03	1			
LnTA	0.54	0.01	0.77	1		
BM	-0.01	0	-0.09	0	1	
PM	-0.01	-0.01	0.05	-0.02	0	1

Panel F: Multicollinearity Test ENV. SOC and GOV Specification

Correlation	ENV	SOC	GOV	RMRF	LMC	LTA	BM	PM
ENV	1							
SOC	0.73	1						
GOV	0.39	0.43	1					
RMRF	0.04	0.04	0.01	1				
LnMC	0.37	0.41	0.33	0.03	1			
LnTA	0.5	0.47	0.37	0.01	0.77	1		
BM	-0.01	-0.02	0.01	0	-0.09	0	1	
PM	-0.01	-0.02	0.02	-0.01	0.05	-0.02	0	1

Item 7

The difference in return between leading (high) and lagging (low), HML, sustainability corporations measured by Jensen's alpha. The number within parenthesis is showing the corresponding p-value. The regressions are based on the sustainability dimension, ESG, Env, Soc and Gov, and the model specification, CAPM, FF3FM, C4FM and FF5FM. The regressions are based on 192 observations from July 2002 through June 2018.

Jensen's Alpha	FF5FM	C4FM	FF3FM	CAPM
HMLESG	-0.000	-0.000	-0.000	-0.000
	(0.950)	(0.958)	(0.718)	(0.889)
HMLENV	0.000	-0.000	-0.001	-0.000
	(0.709)	(0.782)	(0.510)	(0.739)
HMLSOC	0.001	0.001	0.000	0.001
	(0.375)	(0.626)	(0.780)	(0.703)
HMLGOV	0.002*	0.001	0.000	0.000
	(0.071)	(0.431)	(0.641)	(0.726)

*** p<0.01, ** p<0.05, * p<0.1

Item 8

Regressions for 6 value-weighted Size-ESG, Size-Env, Size-Soc and Size-Gov portfolios, based on two different model specifications: FF3FM and C4FM. The regressions are based on 192 observations from July 2002 through June 2018. The table is presented to show the parameters related to the FF3FM and C4FM.

	ESG			***ENV***			***SOC***			***GOV***		
	Lagging	*Mid*	*Leading*	*Lagging*	*Mid*	*Leading*	*Lagging*	*Mid*	*Leading*	*Lagging*	*Mid*	*Leading*
Panel A: Fama-French Three Factor Model: Intercept, SMB and HML												
	α			α			α			α		
Small	0.0029*	0.0039**	0.0008	0.00328**	0.00368**	0.00156	0.00210	0.00475**	0.00150	0.00335**	0.00305*	0.00249
Big	0.0022	0.0023	0.0036**	0.00281	0.00260	0.00309*	0.00216	0.00234	0.00333**	0.00181	0.00122	0.00354**
	SMB			SMB			SMB			SMB		
Small	0.338***	0.361***	0.314***	0.330***	0.355***	0.353***	0.351***	0.349***	0.296***	0.328***	0.335***	0.379***
Big	0.263***	0.0783*	-0.0375	0.158***	0.0737*	-0.0348	0.222***	0.0168	-0.0236	0.178***	0.00367	-0.0263
	HML			HML			HML			HML		
Small	-0.153	0.125	0.279**	-0.120	0.0604	0.270**	-0.142	0.0661	0.299**	-0.0563	0.0710	0.0147
Big	-0.326**	-0.0226	0.0803	-0.152**	-0.0545	0.139	-0.367***	0.156**	0.0722	-0.0455	0.125	0.0823
Panel B: Carhart 4 Factor Model: Intercept, SMB, HML and UMD												
	α			α			α			α		
Small	0.0028*	0.0042**	0.00095	0.00305*	0.00396**	0.00202	0.00200	0.00505**	0.00166	0.00331**	0.00313**	0.00289
Big	0.002	0.0022	0.0038**	0.00277	0.00270	0.0032**	0.00218	0.00222	0.00354**	0.00184	0.00119	0.00369**
	SMB			SMB			SMB			SMB		
Small	0.346***	0.343***	0.309***	0.342***	0.340***	0.328***	0.357***	0.333***	0.287***	0.331***	0.331***	0.357***
Big	0.273***	0.0810**	-0.0485	0.160***	0.0686**	-0.0409	0.220***	0.0233	-0.0349	0.176***	0.00525	-0.0342
	HML			HML			HML			HML		
Small	-0.137	0.0897	0.269*	-0.0959	0.0310	0.222*	-0.131	0.0346	0.282**	-0.0518	0.0629	-0.0273
Big	-0.307**	-0.0174	0.0588	-0.148*	-0.0646	0.127	-0.369***	0.168**	0.0499	-0.0494	0.128	0.0668
	UMD			UMD			UMD			UMD		
Small	0.0482	-0.112	-0.0337	0.0750	-0.0932	-0.155	0.0334	-0.0996	-0.0542	0.0143	-0.0257	-0.133
Big	0.0605	0.0165	-0.0682	0.0137	-0.0321	-0.0381	-0.00752	0.0402	-0.0706	-0.0122	0.00980	-0.0493

**** p<0.01, ** p<0.05, * p<0.1*

Item 9

*Present the performance of CAPM, FF3FM, C4FM and FF5FM specifications dependent on the 2*3 sorted Size-ESG, Size-Env, Size-Soc and Size-Gov portfolios. The model specification performance is measured by the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) in the period of July 2002 through June 2018.*

	AIC					BIC				
ESG	FF5F5	C4FM	FF3FM	CAPM	Average	FF5F5	C4FM	FF3FM	CAPM	Average
BLE	-955.3	-933.8	-934.3	-935.3	-939.7	-935.8	-917.5	-921.3	-928.8	-925.8
BM	-897.8	-893.9	-895.8	-892.8	-895.1	-878.3	-877.6	-882.8	-886.3	-881.2
BLA	-870.9	-854.1	-855.3	-796.3	-844.1	-851.4	-837.8	-842.2	-789.8	-830.3
SLE	-861.3	-833.5	-835.3	-758.6	-822.2	-841.8	-817.3	-822.3	-752.1	-808.4
SM	-923.0	-913.8	-912.1	-792.2	-885.3	-903.5	-897.5	-899.1	-785.7	-871.4
SLA	-941.3	-916.2	-917.5	-810.6	-896.4	-921.8	-899.9	-904.4	-804.1	-882.5
Average	-908.3	-890.9	-891.7	-831.0		-888.7	-874.6	-878.7	-824.5	
ENV	FF5F5	C4FM	FF3FM	CAPM	Average	FF5F5	C4FM	FF3FM	CAPM	Average
BLE	-948.8	-933.4	-934.9	-933.6	-937.7	-929.2	-917.1	-921.9	-927.1	-923.8
BM	-932.8	-906.8	-908.5	-905.7	-913.4	-913.2	-890.5	-895.5	-899.1	-899.6
BLA	-886.4	-885.8	-887.7	-864.0	-881.0	-866.9	-869.5	-874.7	-857.5	-867.1
SLE	-840.8	-820.3	-818.0	-734.0	-803.3	-821.2	-804.0	-804.9	-727.5	-789.4
SM	-949.3	-937.1	-936.2	-810.0	-908.2	-929.8	-920.8	-923.2	-803.5	-894.3
SLA	-943.1	-920.4	-920.7	-817.0	-900.3	-923.5	-904.1	-907.7	-810.4	-886.4
Average	-916.9	-900.6	-901.0	-844.0		-897.3	-884.3	-888.0	-837.5	
SOC	FF5F5	C4FM	FF3FM	CAPM	Average	FF5F5	C4FM	FF3FM	CAPM	Average
BLE	-948.9	-925.7	-926.2	-928.5	-932.3	-929.3	-909.5	-913.2	-922.0	-918.5
BM	-918.5	-914.4	-915.9	-914.7	-915.9	-898.9	-898.1	-902.9	-908.2	-902.0
BLA	-866.0	-850.8	-852.8	-805.3	-843.7	-846.4	-834.5	-839.8	-798.7	-829.9
SLE	-855.5	-832.4	-833.8	-763.1	-821.2	-835.9	-816.1	-820.8	-756.5	-807.3
SM	-949.9	-933.7	-932.4	-811.1	-906.8	-930.3	-917.4	-919.4	-804.6	-892.9
SLA	-948.1	-925.5	-927.1	-809.5	-902.6	-928.6	-909.2	-914.1	-803.0	-888.7
Average	-914.5	-897.1	-898.1	-838.7		-894.9	-880.8	-885.0	-832.2	
GOV	FF5F5	C4FM	FF3FM	CAPM	Average	FF5F5	C4FM	FF3FM	CAPM	Average
BLE	-949.5	-926.3	-927.5	-929.4	-933.2	-930.0	-910.0	-914.5	-922.9	-919.3
BM	-933.1	-929.4	-931.4	-932.1	-931.5	-913.6	-913.1	-918.3	-925.6	-917.6
BLA	-891.2	-880.0	-882.0	-854.3	-876.9	-871.7	-863.7	-869.0	-847.8	-863.0
SLE	-879.4	-847.4	-845.7	-750.3	-830.7	-859.8	-831.1	-832.7	-743.8	-816.8
SM	-948.8	-933.6	-935.4	-819.7	-909.4	-929.2	-917.3	-922.3	-813.2	-895.5
SLA	-933.7	-919.1	-921.1	-818.0	-898.0	-914.2	-902.9	-908.1	-811.5	-884.2
Average	-922.6	-906.0	-907.2	-850.7		-903.1	-889.7	-894.1	-844.1	

Item 10

Regressions for 6 value weighted Size-ESG, Size-Env, Size-Soc and Size-Gov portfolios, based on four different model specifications: CAPM, FF3FM, C4FM and FF5FM. The regressions are based on 74 observations from July 2002 through July 2008. FF5FM is the model with most factors and is defined as: $R_{pt} - R_{ft} = \alpha_{pt} + \beta_{pt}(R_{mt} - R_{ft}) + s_{pt}SMB_t + h_{pt}HML_t + r_{pt}RMW_t + c_{pt}CMA_t + e_{pt}$

	ESG			*ENV*			*SOC*			*GOV*		
	Lagging	*Mid*	*Leading*	*Lagging*	*Mid*	*Leading*	*Lagging*	*Mid*	*Leading*	*Lagging*	*Mid*	*Leading*
Panel A: CAPM intercepts: Rm-Rf												
		α			α			α			α	
Small	0.00893**	0.00913**	0.00987**	0.00843**	0.0102**	0.00913*	0.00636	0.0111***	0.0106***	0.00835**	0.00985**	0.00875*
Big	0.00540	0.00490	0.0072***	0.00510	0.00514	0.00659**	0.00583	0.00593**	0.00654**	0.00511	0.00459*	0.00742**
Panel B: Three factor model intercepts: Rm-Rf, SMB and HML												
		α			α			α			α	
Small	0.0088***	0.00606**	0.0087***	0.00797**	0.00784**	0.00667*	0.00577**	0.00861**	0.0091***	0.0071***	0.00776**	0.00785**
Big	0.00667**	0.00526*	0.0084***	0.00536*	0.00629**	0.00742**	0.00676**	0.00633**	0.0075***	0.00514*	0.00500**	0.0085***
Panel C: Four factor model intercepts: Rm-Rf, SMB, HML, UMD												
		α			α			α			α	
Small	0.0089***	0.00676**	0.0089***	0.00776**	0.00843**	0.00774**	0.00589**	0.009***	0.0094***	0.00717**	0.00819**	0.0086***
Big	0.00721**	0.00598**	0.0087***	0.00596**	0.00662**	0.00780**	0.00734**	0.00699**	0.0079***	0.00569*	0.00521**	0.0089***
Panel D: Five factor model coefficients: Rm-Rf, SMB, HML, RMW and CMA												
		α			α			α			α	
Small	0.0082***	0.00531**	0.009***	0.00755**	0.00759**	0.00583*	0.00502**	0.0083***	0.0098***	0.00663**	0.00727**	0.00741**
Big	0.00610**	0.00463	0.0072***	0.00565*	0.00507*	0.00676**	0.00605**	0.00494*	0.0069***	0.00451	0.00499*	0.0072***
		β			β			β			β	
Small	0.810***	0.751***	0.836***	0.745***	0.789***	0.833***	0.788***	0.799***	0.768***	0.793***	0.789***	0.831***
Big	0.850***	0.788***	0.771***	0.894***	0.787***	0.782***	0.824***	0.863***	0.716***	0.861***	0.767***	0.803***
		s			*s*			*s*			*s*	
Small	0.303***	0.364***	0.304***	0.303***	0.367***	0.327***	0.330***	0.345***	0.296***	0.315***	0.326***	0.368***
Big	0.243***	0.0733	-0.0620*	0.146***	0.0437	-0.0465	0.217***	-0.0463	-0.0295	0.188***	0.0349	-0.0741*
		h			*h*			*h*			*h*	
Small	-0.0305	0.436*	-0.0457	0.00404	0.128	0.496*	0.00564	0.331	-0.134	0.0933	0.203	0.0944
Big	-0.160	0.145	0.323	-0.206	0.271	0.230	0.0513	0.391*	0.252	0.0645	-0.0971	0.438**
		r			*r*			*r*			*r*	
Small	-0.384***	-0.363**	-0.419***	-0.423***	-0.286*	-0.509***	-0.329**	-0.474***	-0.331**	-0.342**	-0.292**	-0.576***
Big	-0.473***	-0.270	-0.271**	-0.197	-0.407***	-0.247**	-0.588***	-0.107	-0.396***	-0.315*	-0.0467	-0.358***
		c			*c*			*c*			*c*	
Small	-0.416**	-0.438*	-0.125	-0.371*	-0.243	-0.551**	-0.421*	-0.368*	0.0352	-0.343	-0.317	-0.470**
Big	-0.450*	-0.349	-0.535***	-0.0256	-0.611**	-0.346*	-0.558**	-0.489**	-0.405*	-0.376	-0.0300	-0.604***

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Item 11

Regressions for 6 value weighted Size-ESG, Size-Env, Size-Soc and Size-Gov portfolios, based on four different model specifications: CAPM, FF3FM, C4FM and FF5FM. The regressions are based on 111 observations from April 2009 through June 2018. FF5FM is the model with most factors and is defined as: $R_{pt} - R_{ft} = \alpha_{pt} + \beta_{pt}(R_{mt} - R_{ft}) + s_{pt}SMB_t + h_{pt}HML_t + r_{pt}RMW_t + c_{pt}CMA_t + e_{pt}$

	ESG			ENV			SOC			GOV		
	Lagging	*Mid*	*Leading*	*Lagging*	*Mid*	*Leading*	*Lagging*	*Mid*	*Leading*	*Lagging*	*Mid*	*Leading*
Panel A: CAPM intercepts: Rm-Rf												
		α			α			α			α	
Small	0.00113	0.00369	-0.00165	0.00219	0.00258	0.000334	0.00138	0.00407*	-0.00156	0.00232	0.00144	0.00209
Big	0.00193	0.00170	0.00109	0.00236	0.00152	0.000673	0.00134	0.00108	0.00105	0.00163	-0.000235	0.000905
Panel B: Three factor model intercepts: Rm-Rf, SMB and HML												
		α			α			α			α	
Small	0.000530	0.00341*	-0.00156	0.00162	0.00221	0.000423	0.000761	0.00378**	-0.00146	0.00185	0.00101	0.00192
Big	0.00149	0.00175	0.00166	0.00195	0.00179	0.00125	0.000907	0.00132	0.00162	0.00166	0.000174	0.00139
Panel C: Four factor model intercepts: Rm-Rf, SMB, HML, UMD												
		α			α			α			α	
Small	0.000299	0.00337*	-0.00147	0.00140	0.00219	0.000494	0.000567	0.00376**	-0.00127	0.00168	0.000866	0.00203
Big	0.000863	0.00138	0.00176	0.00158	0.00168	0.00123	0.000444	0.000968	0.00168	0.00131	5.37e-05	0.00142
Panel D: Five factor model coefficients: Rm-Rf, SMB, HML, RMW and CMA												
		α			α			α			α	
Small	0.00248	0.00541**	0.00186	0.00370	0.00424*	0.00409	0.00338	0.00509**	0.00288	0.00308	0.00343	0.00545**
Big	0.00157	0.00100	0.00441**	0.000990	0.00291	0.00365*	0.000248	0.00111	0.00438**	0.00162	0.00128	0.00436**
		β			β			β			β	
Small	0.845***	0.775***	0.853***	0.827***	0.824***	0.787***	0.873***	0.753***	0.790***	0.821***	0.853***	0.789***
Big	0.803***	0.928***	0.795***	0.865***	0.812***	0.838***	0.889***	0.852***	0.804***	0.866***	0.864***	0.819***
		s			*s*			*s*			*s*	
Small	0.305***	0.278***	0.297***	0.301***	0.279***	0.268***	0.309***	0.250***	0.288***	0.271***	0.297***	0.284***
Big	0.175***	0.0412	-0.0849**	0.150***	-0.0141	-0.0648	0.0880*	0.0619	-0.0935**	0.0924*	-0.0248	-0.0548
		h			*h*			*h*			*h*	
Small	-0.0354	0.00673	0.524***	-0.0574	0.0763	0.439***	-0.0669	0.136	0.423***	0.0211	0.0405	0.169
Big	-0.0732	0.0327	0.121	-0.00215	0.176	0.124	-0.191	0.225*	0.0990	0.130	0.148	0.0871
		r			*r*			*r*			*r*	
Small	-0.221	-0.310**	-0.311	-0.249	-0.234*	-0.348*	-0.283*	-0.154	-0.434**	-0.156	-0.282**	-0.390**
Big	-0.0346	0.0126	-0.302**	0.0944	-0.0906	-0.299**	0.0437	-0.0264	-0.306**	-0.0346	-0.155	-0.333**
		c			*c*			*c*			*c*	
Small	0.0888	0.372**	-0.100	0.139	0.106	-0.0543	0.0748	0.0794	0.00734	0.113	0.138	0.129
Big	0.0901	0.206	0.0933	0.00503	-0.0694	0.201	0.0728	0.158	0.104	0.127	0.148	0.123

**** p<0.01, ** p<0.05, * p<0.1*

Item 12

Panel data fixed effect regressions in the period from July 2002 through July 2008 with four different model specifications. In Table A is the sustainability scores from the period t, in table B is the sustainability scores from t-1, in Table C is the sustainability scores constructed by taking the first difference and Table D is the sustainability scores constructed by taking the lagged first difference. As seen in column 1-5 in table A-D, different combinations of the sustainability scores are used.

Table A

Variable	1A	2A	3A	4A	5A
ESG_t	-0.003				
	(0.005)				
ENV_t		-0.009**	-0.006*		
		(0.004)	(0.004)		
SOC_t		0.006		0.002	
		(0.005)		(0.004)	
GOV_t		-0.003			-0.004
		(0.005)			(0.004)
RMRF	0.328***	0.327***	0.328***	0.327***	0.330***
	(0.013)	(0.014)	(0.013)	(0.013)	(0.013)
BTM	0.003	-0.004	-0.019	-0.017	-0.006
	(0.060)	(0.065)	(0.061)	(0.061)	(0.063)
LnMC	24.739***	24.171***	24.598***	24.926***	24.403***
	(1.307)	(1.392)	(1.308)	(1.326)	(1.309)
LnTA	-3.370***	-3.571***	-3.540***	-3.045***	-3.210***
	(0.746)	(0.769)	(0.742)	(0.787)	(0.764)
PM	-0.001	-0.001	-0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Obs.	25,197	23,106	25,416	25,163	25,383
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

Table B

Variable	1B	2B	3B	4B	5B
ESG_{t-1}	-0.034***				
	(0.011)				
ENV_{t-1}		-0.010	-0.014**		
		(0.008)	(0.007)		
SOC_{t-1}		-0.014		-0.018**	
		(0.009)		(0.008)	
GOV_{t-1}		-0.009			-0.011
		(0.009)			(0.008)
RMRF	0.331***	0.329***	0.334***	0.329***	0.333***
	(0.014)	(0.015)	(0.014)	(0.014)	(0.014)
BTM	0.020	0.026	0.004	0.034	0.033
	(0.089)	(0.097)	(0.093)	(0.096)	(0.094)
LnMC	24.452***	23.997***	24.321***	24.557***	24.187***
	(1.391)	(1.425)	(1.379)	(1.391)	(1.370)
LnTA	-3.347***	-3.456***	-3.298***	-3.245***	-3.393***
	(0.782)	(0.791)	(0.765)	(0.796)	(0.762)
PM	-0.001	-0.002	-0.001	-0.002	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Obs.	22,958	21,511	23,198	22,885	23,152
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

Table C

Variable	1C	2C	3C	4C	5C
ΔESG_t	0.008				
	(0.006)				
ΔENV_t		-0.006	-0.002		
		(0.004)	(0.004)		
ΔSOC_t		0.009*		0.008*	
		(0.005)		(0.005)	
ΔGOV_t		0.001			0.001
		(0.004)			(0.004)
RMRF	0.330***	0.329***	0.334***	0.328***	0.333***
	(0.014)	(0.015)	(0.014)	(0.014)	(0.014)
BTM	0.022	0.029	0.007	0.034	0.035
	(0.089)	(0.097)	(0.093)	(0.096)	(0.093)
LnMC	24.451***	23.999***	24.318***	24.567***	24.180***
	(1.393)	(1.427)	(1.380)	(1.391)	(1.371)
LnTA	-3.386***	-3.504***	-3.318***	-3.266***	-3.410***
	(0.780)	(0.787)	(0.764)	(0.796)	(0.759)
PM	-0.001	-0.002	-0.001	-0.002	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Obs.	22,958	21,511	23,198	22,885	23,152
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

Table D

Variable	1D	2D	3D	4D	5D
ΔESG_{t-1}	-0.017				
	(0.011)				
ΔENV_{t-1}		-0.014**	-0.015**		
		(0.007)	(0.006)		
ΔSOC_{t-1}		-0.004		-0.006	
		(0.009)		(0.008)	
ΔGOV_{t-1}		-0.000			-0.001
		(0.008)			(0.007)
RMRF	0.336***	33.641***	0.338***	0.338***	0.339***
	(0.015)	(1.499)	(0.015)	(0.015)	(0.015)
BTM	-0.018	-0.023	-0.015	-0.020	-0.014
	(0.072)	(0.075)	(0.073)	(0.075)	(0.073)
LnMC	24.714***	24.529***	24.754***	24.850***	24.781***
	(1.525)	(1.554)	(1.527)	(1.536)	(1.524)
LnTA	-3.162***	-2.948***	-2.950***	-2.829***	-2.960***
	(0.785)	(0.803)	(0.779)	(0.802)	(0.802)
PM	-0.001	-0.149	-0.002	-0.002	-0.002
	(0.001)	(0.125)	(0.001)	(0.001)	(0.001)
Obs.	21,73	21,014	21,92	21,655	21,913
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

**** p<0.01, ** p<0.05, * p<0.1, TFE = Time Fixed Effect, FFE = Firm Fixed Effect*

Item 13

Panel data fixed effect regressions in the period from April 2009 through June 2018 with four different model specifications. In Table A is the sustainability scores from the period t, in table B is the sustainability scores from t-1, in Table C is the sustainability scores constructed by taking the first difference and Table D is the sustainability scores constructed by taking the lagged first difference. As seen in column 1-5 in table A-D, different combinations of the sustainability scores are used.

Table A

Variable	1A	2A	3A	4A	5A
ESG_t	-0.072***				
	(0.006)				
ENV_t		-0.019***	-0.043***		
		(0.006)	(0.005)		
SOC_t		-0.047***		-0.054***	
		(0.005)		(0.004)	
GOV_t		-0.007*			-0.017***
		(0.004)			(0.004)
RMRF		0.416***	0.418***	0.417***	0.419***
		(0.013)	(0.011)	(0.012)	(0.011)
BTM		0.001***	0.001***	0.001***	0.001***
		(0.000)	(0.000)	(0.000)	(0.000)
LnMC		22.354***	22.322***	22.087***	22.198***
		(1.126)	(1.053)	(1.048)	(1.018)
LnTA		-1.249	-1.312	-1.160	-1.150
		(0.857)	(0.800)	(0.811)	(0.781)
PM		0.000	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)
Obs.	56.378	48.096	55.991	55.003	55.847
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

Table B

Variable	1B	2B	3B	4B	5B
ESG_{t-1}	-0.088***				
	(0.015)				
ENV_{t-1}		-0.038**	-0.056***		
		(0.016)	(0.013)		
SOC_{t-1}		-0.048***		-0.062***	
		(0.012)		(0.010)	
GOV_{t-1}		-0.010			-0.018*
		(0.009)			(0.010)
RMRF	0.411***	0.412***	0.412***	0.417***	0.417***
	(0.012)	(0.013)	(0.012)	(0.012)	(0.012)
BTM	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LnMC	22.856***	22.488***	22.402***	22.080***	21.874***
	(1.103)	(1.182)	(1.105)	(1.146)	(1.100)
LnTA	-1.605**	-1.521*	-1.453*	-1.468*	-1.594**
	(0.735)	(0.788)	(0.751)	(0.772)	(0.720)
PM	-0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Obs.	51.42	45.313	50.949	49.857	50.813
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

Table C

Variable	1C	2C	3C	4C	5C
ΔESG_t	-0.048***				
	(0.006)				
ΔENV_t		-0.012**	-0.029***		
		(0.006)	(0.005)		
ΔSOC_t		-0.037***		-0.044***	
		(0.005)		(0.005)	
ΔGOV_t		-0.003			-0.010**
		(0.004)			(0.004)
RMRF	0.415***	0.417***	0.414***	0.421***	0.418***
	(0.012)	(0.013)	(0.012)	(0.012)	(0.012)
BTM	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LnMC	22.707***	22.323***	22.310***	21.952***	21.860***
	(1.095)	(1.173)	(1.101)	(1.141)	(1.099)
LnTA	-1.722**	-1.619**	-1.528**	-1.526*	-1.622**
	(0.740)	(0.796)	(0.755)	(0.778)	(0.721)
PM	-0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Obs.	51.42	45.313	50.949	49.857	50.813
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

Table D

Variable	1D	2D	3D	4D	5D
ΔESG_{t-1}	0.019				
	(0.012)				
ΔENV_{t-1}		0.009	0.014		
		(0.014)	(0.011)		
ΔSOC_{t-1}		-0.007		-0.002	
		(0.011)		(0.009)	
ΔGOV_{t-1}		0.012			0.016*
		(0.009)			(0.009)
RMRF	0.423***	42.367***	0.426***	0.426***	0.426***
	(0.013)	(1.352)	(0.013)	(0.013)	(0.013)
BTM	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LnMC	22.551***	22.239***	21.924***	21.863***	21.752***
	(1.138)	(1.187)	(1.141)	(1.177)	(1.127)
LnTA	-1.607**	-1.474*	-1.392*	-1.463*	-1.426*
	(0.782)	(0.831)	(0.797)	(0.820)	(0.777)
PM	0.000	0.023	0.000	0.000	0.000
	(0.000)	(0.023)	(0.000)	(0.000)	(0.000)
Obs.	48.233	44.757	47.797	46.9	47.677
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

**** p<0.01, ** p<0.05, * p<0.1, TFE = Time Fixed Effect, FFE = Firm Fixed Effect*

Item 14

Panel data fixed effect regressions in the period from July 2002 through June 2018 with four different model specifications. In Table A is the sustainability scores from the period t, in table B is the sustainability scores from t-1, in Table C is the sustainability scores constructed by taking the first difference and Table D is the sustainability scores constructed by taking the lagged first difference. As seen in column 1-5 in table A-D, different combinations of the sustainability scores are used.

Table A					
Variable	1A	2A	3A	4A	5A
ESG_t	-0.013***				
	(0.003)				
ENV_t		-0.004	-0.006***		
		(0.003)	(0.002)		
SOC_t		-0.004		-0.006***	
		(0.003)		(0.002)	
GOV_t		-0.005**			-0.007***
		(0.002)			(0.002)
RMRF	0.362***	0.361***	0.363***	0.363***	0.362***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
BTM	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LnMC	22.517***	22.141***	22.256***	22.373***	22.427***
	(0.620)	(0.665)	(0.635)	(0.638)	(0.628)
LnTA	-2.230***	-2.146***	-2.161***	-2.009***	-2.002***
	(0.534)	(0.564)	(0.532)	(0.552)	(0.539)
PM	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Obs.	85,361	74,577	85,216	83,972	85,085
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

Table B					
Variable	1B	2B	3B	4B	5B
ESG_{t-1}	-0.022***				
	(0.008)				
ENV_{t-1}		0.006	-0.002		
		(0.007)	(0.006)		
SOC_{t-1}		-0.023***		-0.016***	
		(0.007)		(0.006)	
GOV_{t-1}		-0.004			-0.005
		(0.006)			(0.006)
RMRF	0.362***	0.360***	0.364***	0.361***	0.362***
	(0.0102)	(0.011)	(0.010)	(0.010)	(0.010)
BTM	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LnMC	22.34***	22.009***	22.149***	22.096***	22.115***
	(0.650)	(0.684)	(0.661)	(0.664)	(0.660)
LnTA	-2.280***	-2.254***	-2.192***	-2.156***	-2.334***
	(0.508)	(0.531)	(0.509)	(0.525)	(0.506)
PM	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Obs.	77,821	70,014	77,628	76,216	77,522
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

Table C					
Variable	1C	2C	3C	4C	5C
ΔESG_t	-0.016***				
	(0.003)				
ΔENV_t		-0.010***	-0.012***		
		(0.003)	(0.002)		
ΔSOC_t		-0.003		-0.008***	
		(0.003)		(0.002)	
ΔGOV_t		-0.004*			-0.005***
		(0.002)			(0.002)
RMRF	0.361***	0.359***	0.364***	0.361***	0.362***
	(0.0102)	(0.011)	(0.010)	(0.010)	(0.010)
BTM	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LnMC	22.33***	22.000***	22.150***	22.095***	22.110***
	(0.650)	(0.683)	(0.661)	(0.664)	(0.660)
LnTA	-2.326***	-2.304***	-2.221***	-2.183***	-2.343***
	(0.509)	(0.534)	(0.511)	(0.526)	(0.507)
PM	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Obs.	77,821	70,014	77,628	76,216	77,522
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

Table C					
Variable	1D	2D	3D	4D	5D
ΔESG_{t-1}	-0.036***				
	(0.009)				
ΔENV_{t-1}		-0.027***	-0.030***		
		(0.007)	(0.006)		
ΔSOC_{t-1}		-0.012		-0.023***	
		(0.008)		(0.007)	
ΔGOV_{t-1}		0.000			-0.001
		(0.006)			(0.006)
RMRF	0.366***	36.254***	0.366***	0.366***	0.366***
	(0.011)	(1.110)	(0.011)	(0.011)	(0.011)
BTM	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LnMC	22.370***	22.115***	22.051***	22.095***	22.091***
	(0.690)	(0.707)	(0.695)	(0.707)	(0.688)
LnTA	-2.033***	-1.947***	-1.932***	-1.851***	-1.948***
	(0.545)	(0.570)	(0.548)	(0.565)	(0.552)
PM	0.000	0.019	0.000	0.000	0.000
	(0.000)	(0.023)	(0.000)	(0.000)	(0.000)
Obs.	73,650	69,344	73,464	72,288	73,417
TFE	Yes	Yes	Yes	Yes	Yes
FFE	Yes	Yes	Yes	Yes	Yes

**** p<0.01, ** p<0.05, * p<0.1, TFE = Time Fixed Effec, FFE = Firm Fixed Effect*